

# **ajer**

**THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH**

**VOLUME XXX**

**NUMBER 1**

**MARCH 1984**

**PUBLISHED BY  
THE UNIVERSITY OF ALBERTA • EDMONTON**



# THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

*A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.*

Published quarterly in March, June, September, December by the  
Faculty of Education, The University of Alberta

## CONSULTING EDITORS

D. C. Bjorkquist  
University of Minnesota

P. E. Bryant  
University of Oxford

B. M. Chung  
Chungbuk University,  
Republic of Korea

A. Côté  
University of Ottawa

D. Fitzgerald  
The University of New England

D. B. Harrison  
The University of Calgary

N. L. Hersom  
The University of Saskatchewan

E. Miklos  
The University of Alberta

P. Robinson  
Atlantic Institute of Education

R. E. Traub  
Ontario Institute for  
Studies in Education

K. Watson  
University of Reading

R. J. Wilson  
Queen's University

EDITOR: C. S. Bumbarger

SECRETARY: A. Onishenko

## FACULTY PUBLICATIONS COMMITTEE

E. Miklos (Chairman)

S. I. Jamieson (Operations Officer)

M. E. Compton (Secretary)

J. H. Balderson

R. J. Carney

A. K. Clark

M. McNay

M. Rodda

M. J. Van Manen

*Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, and *Social Science Citation Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, *Sociology of Education Abstracts*, and *Language Behavior Abstracts*.

The subscription rate is \$12.00 per year; single copies are \$3.50 each. Please make cheques payable to The Alberta Journal of Educational Research. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to the Editor, *The Alberta Journal of Educational Research*, Faculty of Education, 845 Education South, The University of Alberta, Edmonton, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436



ISSN 0002-4805

# The Alberta Journal of Educational Research

*Volume XXX, Number 1*


*March, 1984*

## CONTENTS

|  |    |
|--|----|
| Development and Initial Results of a Longitudinal<br>Secondary Follow-up Study . . . . .                         | 1  |
| <i>B. Levin</i>  |    |
| Characteristics of Learning Disabilities Persisting<br>into Adolescence . . . . .                                | 14 |
| <i>L. A. Whyte</i>   |    |
| Supervising Student Teachers: A Means of Professional Renewal ? . . . . .  | 26 |
| <i>M. F. Wideen and D. Hopkins</i>   |    |
| The Link Between Universities and Employment in Alberta . . . . .  | 38 |
| <i>M. Westbury</i>   |    |
| Second Language Teaching Styles: Making or Breaking Attitudes . . . . .  | 49 |
| <i>G. M. H. Gayle</i>  |    |
| WISC-R Subtest Patterning of Below Average, Average, and<br>Above Average IQ Children: A Meta-Analysis . . . . . | 68 |
| <i>H. H. Mueller, U. N. Dash, D. W. Matheson and R. H. Short</i>   |    |
| LIST OF PUBLICATIONS RECEIVED . . . . .  | 86 |

FACULTY OF EDUCATION  
*The University of Alberta*





Digitized by the Internet Archive  
in 2024 with funding from  
University of Alberta Library

[https://archive.org/details/ajer\\_1984](https://archive.org/details/ajer_1984)



BENJAMIN LEVIN

Manitoba Department of Education

## Development and Initial Results of a Longitudinal Secondary Follow-up Study

*Follow-up studies of secondary school students have the potential to contribute significantly to our understanding of the relationship between secondary school, further education, and employment. However, many of the studies which have been done suffer from serious methodological problems, including poor response rates and weak conceptualization. The Peel Secondary Follow-up Study, involving a multi-cohort, multi-contact design, is described. Results from the first round of the study raise important issues of students' expectations and the way in which these are—or, just as frequently, are not—borne out.*

In theory, at least, schools are justified largely in terms of what will happen to students at some later time. Whereas a business or a hospital attempts to secure desired outcomes immediately—this year's profit or this week's cures—schools have the difficult task of showing that they have a positive impact on people years after these people have left the institution.

The follow-up study is one way of trying to find out something about what happens to students after they leave school and, by implication, a way of making judgments about school programs. The literature on follow-up studies is reviewed in this paper, some of the difficulties in conducting such studies are commented on, and an initial attempt by a large Canadian Board of Education to design and implement a follow-up study which would overcome some of these difficulties is described.

### *Review of Literature*

The intrinsic appeal of the follow-up study idea leads one to suppose that such studies are common. However relatively few have been reported in the literature. A

---

Benjamin Levin is Director of Planning and Research for the Manitoba Department of Education, and was formerly Chief Research Officer with the Peel Board of Education in Mississauga, Ontario. He holds a PhD from the Ontario Institute for Studies in Education. His main interest at present is in the application of data and research to policy making in education.



search of sources including RIE, CIJE, ONTERIS, and the Canadian Education Index revealed fewer than two dozen secondary school follow-up studies conducted since 1972. Most of the references obtained were from RIE and were thus unpublished documents, often written to meet the requirements of various United States federal funding programs.

There is considerable range in the studies in terms of size and sophistication. Most common are school district studies which attempt to contact recent graduates and assess the value of their high school education or some portion of it (Finn, 1977; Preston, 1975; Sprumont, 1977). Some more extensive efforts have been made at a state or provincial level. For example, Anisef, Paasche and Turritin (1980) resurveyed students who had initially been surveyed in 1974 while in Grade 12 in Ontario. A series of studies in Manitoba (Russell, 1980) culminated in a review of the relationship of expressed high school intentions with actual postsecondary activities.

The largest studies, however, have been national studies in the United States which have followed students over a long period of time. Project Talent sampled high school students in 1960, and again one, five, and eleven years later. However, by the final survey less than a third of the sample could be located (Call, Otto & Spenner, 1982, p. 63). The more recent National Longitudinal Study of the Class of 1972 includes data on 20,000 students who were seniors in 1973, and who have been resurveyed in 1974, 1976, and 1979. The NLS has also made its data base, containing more than 3,500 pieces of information about each respondent, available to other researchers. By 1981 some 320 studies had been completed or were in progress using NLS data (Taylor, Stafford & Place, 1981). NLS has now been superseded by the High School and Beyond study, with data on 58,000 high school seniors and sophomores in 1980, plus responses from their parents and teachers.

The flavour of the literature can best be given by discussing several methodological dimensions of the work. These include issues of basic design, data collection methods, comprehensiveness of response, and content.

### *Design*

Most follow-up studies are one-time efforts; that is, a single group of former students is contacted at some point after leaving the school. However, there are variations on this procedure. The first is to include students from more than one cohort in the study. For example, Sprumont (1977) included students from five consecutive graduating classes. Harris, Abrams and Simpson (1972) included two consecutive classes of graduates. Righthand (1977) surveyed graduates from 1965, 1968, and 1971. A second variation, a more rarely used procedure, is to contact the same set of students on more than one occasion after they leave school (Smedley & Olson, 1975). This is the procedure used in the National Longitudinal Study, which has surveyed 1972 high school seniors in 1974, 1976, and 1979. Anisef, Paasche and Turritin (1980) and Russell (1980) also resurveyed students after a period of several years. Both these procedures, but especially the second one, create considerable problems in terms of finding students and keeping track of them for later contact. Perhaps for this reason, in only a few studies is an attempt made to contact students more than two or three years after they leave school (e.g., Russell, 1980; Bettis & Hyder, 1978; Crim & Ross, 1976).



### *Data Collection Methods*

Two basic methods have been used for follow-up studies. Mailed questionnaires are more common, with telephone interviews a second choice. Krucek (1975) compared these methods as well as personal interviews, and found that telephone interviews combined with mail questionnaires provided a higher response rate at lower cost. Similar results at the junior college level were obtained by Sheldon (1977).

The evidence on mailed questionnaires indicates that at least three mailings are required to obtain an adequate response. Harris, Abrams and Simpson (1972) obtained responses from 35% on the first mailing, but 80% after three mailings and a telephone reminder. Preston (1975) had a response rate of 72% after three mailings and two telephone reminders. Krucek (1975) had returns from 50% after three mailings. Lewis (1976) reports a 75% response after mailings. These studies include a variety of other devices, such as personal letters from high school principals, reminder cards, and the like to increase response rates. Of the telephone studies, Crim & Ross (1976) and Krucek (1975) both report a 66% response.

### *Response Rates*

A crucial aspect of any follow-up study is the final response rate. These vary greatly, from as low as 12% (Bettis & Hyder, 1978), or 28% (Finn, 1977) to as high as 80% (Harris, Abrams & Simpson, 1972). For the studies reviewed, a response rate of 50% to 60% was about average, and there is evidence (Righthand, 1977; Smedley & Olson, 1975) that many unpublished studies have even lower returns. On the other hand, the National Longitudinal study has been able to maintain a response rate of some 85%, even seven years after the initial survey.

Response rates vary on a number of dimensions. Not surprisingly, rates are higher when the study is done soon after students leave school. The longer the wait, the lower the response rate in general, although NLS is an exception. There is considerable evidence that respondents and nonrespondents vary on significant dimensions (Taylor, Stafford & Place, 1981; Herrnstadt, Horowitz & Sum, 1979). Furthermore, those in the sample who are not located may well be significantly different from those who are contacted (Call, Otto & Spenner, 1982).

The studies reviewed vary in the intensity of efforts to contact all former students in the sample. In some studies no such efforts were made (Righthand, 1977). In other cases efforts to obtain new addresses for students who had moved were made, with varying success. However, the potential for reaching all former students is illustrated by two methodological studies. Call, Otto and Spenner (1982) attempted in the Career Development Study to find 7,000 Washington State students 13 years after high school. Ninety-six percent were located. Clarridge, Havser and Sheehy (1976) found 97.4% of more than 10,000 former students in Wisconsin, 17 years after they left high school! Clarridge's results were aided by additional data obtained from respondents in 1957, such as information on parents' occupation. These studies illustrate that response rates can be very high if sufficient effort is invested in the study.

### *Content*

There is far less variance among studies in terms of the questions asked of students than in the methodological elements. For the most part data fall into two areas. Much of the information concerns occupational and educational outcomes for



former students—jobs held, earnings, unemployment if any, areas of further education, on-the-job training, and so on. All of the studies have collected at least some of this information. In addition, most studies have asked respondents to comment on various aspects of their high school education. Common areas here include general quality of program, relevance to current activity, adequacy of counselling services, preparation for employment and the like.

To review all the results from previous studies is beyond the scope of this paper. In many cases, however, it must be noted that results are called into question by the low response rates of the studies.

Two issues are of particular interest. As noted above, a few studies have sought to compare students' intentions as expressed while in high school with their actual activity after leaving high school. Russell (1980) found that the reliability of expressed intentions of Grade 12 students in Manitoba was higher for those planning University education (70% of whom actually enrolled) than for those planning community college education (30% of whom enrolled). In a study in Ontario, Anisef, Paasche, and Turritin (1980), found that 78% of those intending university enrolled, 56% of those intending community college enrolled, and 88% of those intending to work did so. In a similar vein, Alexander and Cook (1979) argue that students' expressed educational and career plans are often extemporaneous. Thus, it would appear that institutions should not rely heavily on the expressed intentions of students. Secondly, there is evidence that data on income and occupation provided by respondents are of uncertain validity (Taylor, Stafford & Place, 1981).

### *Design of the Study*

The Peel Secondary Follow-up Study was designed to provide information to the Peel system on the post-high school outcomes of Peel students. Its primary purpose was to assist program planning for secondary schools, particularly in light of some of the major suggestions for redirection which were contained in the report of Ontario's Secondary Education Review Project (SERP). Once the actual post-high school activities of Peel students are better known, it should be possible to improve their fit with school programs. In the first round of the study it was decided to collect only basic data about respondents' current activities—status (worker, student, etc.), place of study or employment, and specific job or program. In addition, respondents would be asked whether their current status was what they had anticipated when in Grade 11. More data, if desired, could be built into the 1984 study.

The study was designed specifically to overcome some of the major flaws in the previous literature. Four of these were of particular concern. Unless a high response rate could be obtained, the results would not safely allow generalization for all Peel students. Thus a key concern was achieving a high response rate. A second weakness in the literature is the very short interval in many studies between students' departure from school and the data collection. It was felt that career paths would not be visible for several years after leaving high school. Thus the study would have to reach students some time after they left school.

Similarly, because people's intentions and activities change over time, a multiple contact approach, in which each respondent is contacted more than once, was called for. Such an approach allows a picture to emerge of changes in individuals' lives. Finally, in order to get a clearer picture of school effects, it was thought necessary



to include more than one cohort in the study. The design would then allow comparison both across years and across cohorts of students.

The final design, then, called for the inclusion in the study of a new cohort of students every two years, and the contacting of each cohort every two years for ten years. The first “round” of the study would include two groups of students, but when fully developed, there would be five groups of students in each round, ranging from two to ten years after their initial identification (see Figure 1).

| Year Sampled | Years for Data Collection |      |      |      |      |
|--------------|---------------------------|------|------|------|------|
|              | 1982                      | 1984 | 1986 | 1988 | 1990 |
| 1977-78      | x                         | x    | x    | x    |      |
| 1979-80      | x                         | x    | x    | x    | x    |
| 1981-82      |                           | x    | x    | x    | x    |
| 1983-84      |                           |      | x    | x    | x    |
| 1985-86      |                           |      |      | x    | x    |
| 1987-88      |                           |      |      |      | x    |

Figure 1. Follow-up study design.

Telephone interviews were chosen as the data collection method for the first round. Because the questions to be asked were relatively non-controversial, anonymity was not felt to be a major concern. Telephone interviewing provides a more personal contact than mailing questionnaires. Furthermore, the evidence from the literature was that telephone interviews would provide a higher response rate at a lower cost than would mailed questionnaires. This was particularly true in a study such as this one, in which the interview takes only two or three minutes to conduct. Interviewers were hired on contract to conduct the interviews. In later rounds of the study, however, a variety of data collection methods will likely be used.

Grade 11 was selected as the base year for sampling purposes. The great majority of students are still at school at this age, so that the sample would be similar to the entering high school population. Further, grade 11 students are able to give informed consent to participate in the study. As well, the time between initial contact in Grade 11 and the first survey, two years later, would mean fewer students “lost” due to moves.

For the 1981 Grade 11 sample, as will be done in future samples, students were contacted while in Grade 11 and asked to agree to participate in the study, whose purpose and ongoing nature were explained. Students were also given the option of refusing to take part. Students who agreed were asked to complete an “Agreement to Participate” and to supply information to help locate them should they move—parents’ occupations and employers, and the address of a close friend or relative. These are among the data suggested by Call, Otto and Spenner (1982) as most useful in later locating those in a follow-up study. Students were contacted repeatedly by letter and telephone until they indicated agreement or refusal to take part in the study.

Had we begun with the 1981 sample, however, no data would have been available until 1984, and even then many students would be either still in high school or just out of school. We therefore decided to include cohorts from the 1977-78 and 1979-80 Grade 11 classes. A 10% random sample of Grade 11 students from



each year was drawn from the records of each secondary school. Approximately 750 names were in each sample (see Figure 2).

The remainder of this paper presents data for the first round of the study, including the 1977 and 1979 cohorts.

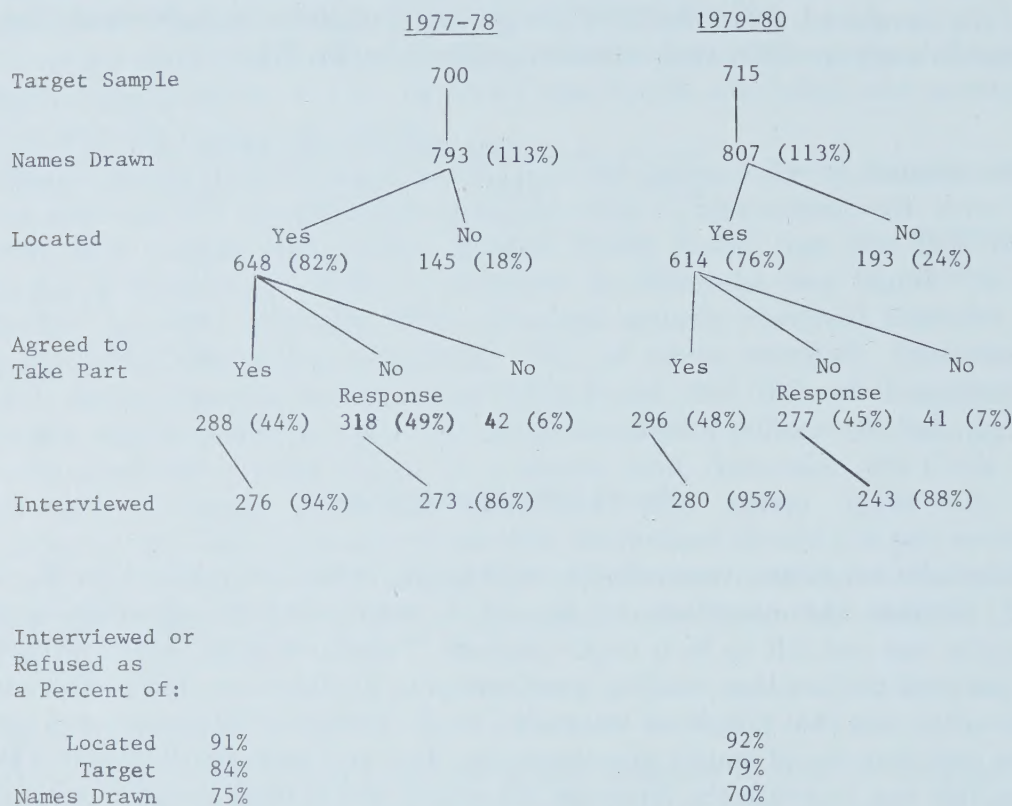


Figure 2. Response rates.

Locating and Contacting Former Students

The first task was to obtain current addresses and telephone numbers for the students in the sample, beginning with the addresses contained in school records. We verified as many of these as we could by checking with current telephone directories. About three-quarters of the families were still at the same addresses or had the same telephone numbers. Where a listing could not be verified in this way, we called the student's previous number in an attempt to obtain a current number. This procedure produced what we thought were accurate addresses for 82% of the 1977 and 76% of the 1979 sample. Of these, 37% of the 1977 cohort and 20% of the 1979 group were changed from the original school records.

Letters explaining the study, and including the "Agreement to Participate," the "Information Sheet," and a stamped return envelope were then sent to located addresses. Word processing equipment was used to produce personal letters and envelopes for each respondent. This mailing produced 44% to 48% completed agreements, 3% refusals, and 3% wrong addresses.

We then began data collection through telephone interviewing. Because the response rate to the letter was only 48%, just about what the literature would lead one to expect (Call, Otto & Spenner, 1982), we decided to try contacting all those in the sample who had not actually returned a refusal form. The interviewers were given separate instructions and procedures for use with nonrespondents.



Contact by telephone proved quite successful. At the time of writing about 95% of those who agreed to participate have provided data. As well, about 90% of the nonrespondents have also been contacted, some of whom (about 8%) refused to participate. Some persons have not yet been reached. A number of the nonrespondents were no longer at the addresses we had for them.

Figure 2 summarizes the results of this section. Of the names in the original sample, at the time of writing this we had data from more than 70%, refusals from 6%, and were still trying to obtain addresses or to contact the remaining 24%.

### *Maintaining High Response Rates*

We concentrated initially on obtaining data from those for whom addresses or phone numbers have been obtained. Only limited efforts have been made to locate other members of the sample. However, even these limited efforts, relying largely on contacting neighbours at the previous address and calling potential relatives drawn from city directories, have been productive. Over the course of the study more than 100 sample members have been located using these techniques.

We are now making more intensive efforts to locate the persons in the sample for whom we do not yet have addresses. Our procedures include, in addition to those mentioned, going back to school records for additional information, looking for siblings now in the schools, and contacting former classmates. We expect, based on our experience so far, to locate at least two-thirds of these remaining sample members. Including data from them and from those not yet contacted, we expect eventually to have data from more than 90% of the original sample. This response rate would be higher than any other documented study done by a school district, and would be exceeded only by such major efforts as the National Longitudinal Study, and the Career Development Study cited earlier.

Although these procedures to increase response rate may seem very troublesome, our data indicate that they are necessary. We compared respondents who returned the Agreement form with those who did not but who provided data when contacted. The two groups differed significantly on several dimensions. Compared with voluntary respondents, nonrespondents were significantly more likely to be male ( $F=24.26$ ,  $df=2$ ,  $p<.001$ ), to be working rather than studying ( $F=21.44$ ,  $df=10$ ,  $p<.02$ ), and not to be from vocational schools ( $F=13.8$ ,  $df=2$ ,  $p<.001$ ).

Fifty-six percent of the 1977 cohort and 48% of the 1979 cohort reported being in the labour force. Of these, 7% and 14% respectively reported being unemployed when surveyed.

Almost all the remaining respondents were students, including 43% of the 1977 group and 52% of the 1979 group. In the former cohort 62% of the student group were in university, 36% in college, and 2% in other educational programs. Not surprisingly, the 1979 group was distributed quite differently, with 80% of the students in grade 13, 16% in colleges, and 4% elsewhere. Most students in both groups—more than 90%—were enrolled in educational institutions near home.

There were no statistically significant differences between males and females in terms of labour force participation or continued education.

Table 1 provides further information on the occupations of the former students. The table shows that most respondents are in the entry-level jobs, such as clerical, labour, or sales. While these three categories account for 55% of the 1977 cohort in the labour force and 47% of the 1979 cohort, they comprise only 38% of the

Ontario labour force as a whole. Not surprisingly, there are very few former students in professional or managerial categories.

TABLE 1  
OCCUPATIONS OF RESPONDENTS

|                    | Clerical | Sales | Service | Production | Labour | Mechanical | Other |
|--------------------|----------|-------|---------|------------|--------|------------|-------|
| 1977-78<br>(N=305) | 33%      | 12%   | 10%     | 8%         | 6%     | 6%         | 25%   |
| 1979-80<br>(N=231) | 22%      | 16%   | 9%      | 12%        | 10%    | 4%         | 27%   |

Examples of the types of jobs under these categories are:

Clerical - secretary, accounts payable clerk, bank teller, order desk clerk.

Sales - cashier, salesperson, pricing department worker.

Service - bartender, travel agent, waiter, security officer, police cadet.

Production - assembly operator, equipment operator, machine operator.

Labour - air conditioning helper, gas attendant, parks maintenance worker.

Mechanical - diesel mechanic, plumber, refrigeration apprentice, telephone maintenance person.

The use of broad categories should not, however, obscure the very great range of actual jobs being held by former students. These range across the entire gamut of occupations in the economy, including all of the major census occupational categories.

Table 2 indicates the length of time respondents report having worked for their current employers. Only a small proportion of students (5%, or 29 in 1977 and 3%, or 15 in 1979) report involvement in a formal job training program such as an apprenticeship. An overwhelming majority of these students are male (86% in 1977 and 73% in 1979). Moreover, many of those in such programs (48% in 1977, 33% in 1979) expected in grade 11 to be doing something else, or had no definite plans.

The Government of Ontario estimates (Ontario Statistics, 1981) that 14% of the appropriate age groups attend University and 11% attend some other form of postsecondary institution. These figures are very similar to those obtained by averaging the results for both Peel cohorts. The proportions of Peel students going on to various postsecondary institutions are similar to those in the Province as a whole.

*Expectations and Reality*

A major focus of the study was to compare respondents' current activity with their high school intentions. Effective programming, career counselling, and preparation in school relies on students' having some reasonable sense of what they might do after high school. All respondents were asked whether their current activity was what they had expected, while in grade 11, to be doing. Responses are shown in Table 4.



TABLE 2  
LENGTH OF STAY WITH PRESENT EMPLOYER

|                 | 0-6 Months | 7-12 Months | 1-2 Years | More Than 2 Years |
|-----------------|------------|-------------|-----------|-------------------|
| 1977-78 (N=264) | 29%        | 26%         | 22%       | 23%               |
| 1979-80 (N=215) | 53%        | 30%         | 11%       | 6%                |

TABLE 3  
EDUCATIONAL ACTIVITIES OF RESPONDENTS

|                               | High School | College | University | Other | TOTAL |
|-------------------------------|-------------|---------|------------|-------|-------|
| 1977-78 % of Students (N=224) | 1%          | 36%     | 62%        | 1%    | 100%  |
| % of Respondents (N=549)      | .5%         | 15%     | 25%        | .5%   | 41%   |
| 1979-80 % of Students (N=283) | 80%         | 16%     | 3%         | 1%    | 100%  |
| % of Respondents (N=523)      | 42%         | 10%     | 1%         | 1%    | 54%   |

TABLE 4  
ACTUAL AND EXPECTED ACTIVITIES

| Current/Expected               | Same Job or Course | Other Job | Other Course | Did Not Know |
|--------------------------------|--------------------|-----------|--------------|--------------|
| <u>1977</u>                    |                    |           |              |              |
| Work (N=291)                   | 28%                | 23%       | 12%          | 37%          |
| School (N=224)                 | 54%                | 12%       | 16%          | 18%          |
| University Students (N=137)    | 61%                | 8%        | 15%          | 16%          |
| Total Cohort (N=527)           | 40%                | 17%       | 15%          | 16%          |
| <u>1979</u>                    |                    |           |              |              |
| Work (N=202)                   | 29%                | 29%       | 18%          | 24%          |
| School (N=283)                 | 81%                | 6%        | 4%           | 9%           |
| Post-Secondary Students (N=59) | 58%                | 12%       | 8%           | 23%          |
| Total Cohort (N=509)           | 58%                | 17%       | 10%          | 15%          |

Those in postsecondary institutions are twice as likely as those working to be fulfilling their grade 11 plans. However, even many continuing students find themselves doing something completely unanticipated as little as two years previously (for the 1979 cohort). In other words, students' intentions are very frequently not borne out in practice. From these data, shown in Table 5, it is possible to infer the grade 11 intentions of all students.

TABLE 5  
GRADE 11 INTENTIONS AND CURRENT ACTIVITY

|                         | 1977 (N=516) | 1979 (N=501) |
|-------------------------|--------------|--------------|
| Intended to work        | 161 (31%)    | 137 (27%)    |
| Working                 | 134 (83%)    | 108 (80%)    |
| Working at intended job | 74 (46%)     | 52 (38%)     |
| At school               | 26 (16%)     | 13 (10%)     |
| Unemployed              | 1 ( 1%)      | 11 ( 8%)     |
| Other                   | 0 -          | 2 ( 2%)      |
| Intended School         | 193 (37%)    | 277 (55%)    |
| At school               | 155 (80%)    | 240 (87%)    |
| Intended course         | 120 (62%)    | 231 (83%)    |
| Working                 | 36 (19%)     | 34 (12%)     |
| Unemployed              | 2 ( 1%)      | 2 ( 1%)      |
| Other                   | 0 -          | 1 -          |
| Intended Job Training   | 17 ( 3%)     | 14 ( 3%)     |
| Job training            | 14 (82%)     | 10 (71%)     |
| Working                 | 3 (18%)      | 3 (22%)      |
| At school               | 0 -          | 1 ( 7%)      |
| Unemployed              | 0 -          | 0 -          |
| Other                   | 0 -          | 0 -          |
| Did Not Know            | 145 (28%)    | 76 (15%)     |
| Working                 | 99 (68%)     | 43 (56%)     |
| At school               | 40 (28%)     | 26 (34%)     |
| Unemployed              | 5 ( 3%)      | 7 ( 9%)      |
| Other                   | 1 ( 1%)      | 0 -          |

Of those students who intended to be working, 80% are working, although usually in a different job than they had hoped. Of those who intended to continue their education, 80% are doing so. Again, however, many of these at the postsecondary level are in a different program than intended. Taken as a whole, the responses indicate that former students' short-term educational plans (i.e., to finish grade 13, or to enter a particular postsecondary program) are carried out reasonably often. However, plans related to work are carried out only in a minority of cases. Comparison of the planned and actual jobs of respondents shows that changes of plans are usually from those of a skilled or higher status job to an actual low skill, low status job (See Table 6).



TABLE 6  
INTENDED AND ACTUAL JOBS

| 1977 Cohort           |                    | 1979 Cohort        |                    |
|-----------------------|--------------------|--------------------|--------------------|
| Intended              | Actual             | Intended           | Actual             |
| Chef                  | Lathe Operator     | Radio/TV           | Shipper/Receiver   |
| Social Worker         | Travel Agent       | Pilot              | Material Handler   |
| College Student       | Truck Driver       | Fashion Designer   | Dry Cleaning Clerk |
| Pilot                 | Civil Technology   | Travel Agent       | Clerk              |
| Veterinarian          | Cake Decorator     | Teacher            | Retail Sales       |
| Fashion Merchandising | Assembly Inspector | Tool and Dye Maker | Truck Driver       |
| Architect             | Cook               | Accounting         | Cashier            |
| Pilot                 | Letter Carrier     | Music Teacher      | Secretary          |
| Model                 | Drug Store Clerk   | Funeral Director   | Gas Cashier        |
| Artist                | Gas Attendant      | Police Officer     | Insurance Agent    |
| Probation Officer     | Accounting Clerk   | Golfer             | Sales Clerk        |
| Hairdresser           | Clerk              | Social Worker      | Clerk              |

*Implications of the Study*

The results of this first round of the study are currently under discussion by various staff members within the Peel Board of Education in terms of their implications for secondary schools in the district. There are a number of ways, not all of them described in this paper, in which the results are of importance locally.

The study has, however, a number of implications for secondary education and, more generally, for research; some of these are particularly noteworthy. Although these are initial results only, they are of sufficient significance to suggest they might be stated as implications to be refined and modified in concurrence with data from later rounds.

Students' postsecondary activities are not predictable at the level of the individual. Since school activities such as counselling and streaming are intended to be predictive for individuals, these findings are of crucial importance. Large numbers of students change their plans shortly after high school, especially those entering the work force.

Such a finding suggests that the schools ought to do less to help students make up their minds and more to assist them in being able to accommodate changes in plans as they occur. That is, the results point not to the need for more help in having students pick careers in high school, but to emphasizing the degree to which change can and will occur, and assisting students in planning for a variety of contingencies. Such an approach would take a major change in attitudes towards the counselling function in high schools.

At the same time, schools need to consider structural changes which will make it easier for students to change their minds—to reenter high school, to change track, to pick up additional courses, or to meet entrance requirements for different kinds of postsecondary education. This sort of programmatic flexibility could be of great assistance to students.

Students also need much more knowledge—experiential knowledge—about what work is like. It is apparent that many students have limited knowledge and very unrealistic ideas about what they will be able to do after high school. While the school ought not to discourage high expectations, it does have a responsibility to tell students about such matters as the actual composition of the labour force and the kinds of entry-level jobs which will likely be available to them. It does have a responsibility to inform them about the real possibilities for entry to apprenticeships or skilled trades. Such programs as cooperative education and work experience would seem to be a clear step in the right direction.

Finally, the early results of the study indicate the value of follow-up studies, especially if they can be made longitudinal. As the Peel study continues, a picture will emerge of the impact of education on school to work transitions, and of the relationship between students' high school plans and experiences and their lives years later. Such a picture may prove to be very humbling; at the same time, it cannot but be very informative.

#### References

- Alexander, K., & Cook, M. (1979). The motivational relevance of educational plans. *Social Psychology Quarterly*, 42, 202-213.
- Anisef, P., Paasche, G., & Turritin, A. (1980). *Is the die cast?* Toronto: Ministry of Colleges and Universities.
- Bettis, G., & Hyder, C. (1978). *A study to develop and implement a one- and five-year student and employer based follow-up study of secondary vocational education programs*. Johnson City, TN: East Tennessee State University. ERIC Document #ED 176 032.
- Call, V., Otto, L., & Spenner, K. (1982). *Tracking respondents: A multi-method approach*. Lexington, MA: D. C. Heath.
- Clarridge, B., Hauser, T., & Sheehy, L. (1976). *Tracing members of a panel: A 17-year follow-up*. Madison, WI: University of Wisconsin. ERIC Document #ED 185 108.
- Crim, R., & Ross, E. (1976). *A follow-up survey of graduates of the high schools in vocational region No. 8, New Hampshire*. ERIC Document #ED 123 471.
- Finn, M. (1977). Follow-up study: Wappinger School District. ERIC Document #ED 158 193.
- Harris, W., Abrams, C., & Simpson, H. (1972). *Project TRACE: Research and evaluation of follow-up studies of high school graduates*. Goleta, CA: Human Factors Research. ERIC Document #ED 079 392.
- Herrnstadt, I., Horowitz, M., & Sum, A. (1979). *Transition from school to work*. Boston, MA: Northeastern University. ERIC Document #ED 183 721.
- Krueck, T. (1975, April). *A comparison of three methods of conducting a follow-up study*. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C. ERIC Document #ED 109 339.
- Lewis, R. (1976). *Graduate follow-up: TEX-SIS follow-up*. Ft. Worth, TX: Texas Education Agency. ERIC Document #ED 128 045.
- Preston, J. (1975). *Follow-up study of Sarasota County's 1974 senior class*. Sarasota, FA: Sarasota Board of Public Instruction. ERIC Document #ED 121 996.



- Preston, J. (1973). *Vocational-technical and adult education: 1974 student and employer follow-up surveys*. Sarasota, FA: Sarasota Board of Public Instruction. ERIC Document #ED 121 997.
- Righthand, H. (1977). *A follow-up study of Connecticut Distributive education students*. Hartford: Connecticut State Department of Education. ERIC Document #ED 141 369.
- Russell, N. (1980). *Post high school outcomes of Manitoba high school students*. Winnipeg: Manitoba Department of Education. ERIC Document #ED 180 401.
- Sheldon, M. S. (1977). *Telephone interview vs. mass mailing for student follow-up*. Los Angeles, CA: Pierce College. ERIC Document #ED 192 848.
- Smedley, R., & Olson, G. (1975). *Graduate follow-up studies: How useful are they?* ERIC Document #ED 109 431.
- Sprumont, B. (1977). *The technical graduate of Scarborough: A survey of employment and further education*. Toronto: Scarborough Board of Education.
- Taylor, M. E., Stafford, C., & Place, C. (1981). *National longitudinal study of the class of 1972 study reports update*. Research Triangle Park, NC: Research Triangle Institute.

LILLIAN A. WHYTE

*The University of Alberta*

## Characteristics of Learning Disabilities Persisting into Adolescence

*A self-rating behavior checklist was administered to learning disabled and academically achieving adolescents. Items involving visual-spatial development, language and communication skills, memory and sequencing, social and emotional adjustment, and attentional skills differentiated between the two groups of students. The results have implications for identifying, educating, and counselling learning disabled adolescents.*

Research concerning the nature of learning disabilities in adolescents is limited at the present time. Anderson (1970) compared the common themes that emerge from anecdotal records (one major source of data) of adolescents who have coped successfully with learning disabilities and those who have not. Anderson concluded that three factors emerge as characteristic of persons who have made a negative adaptation to their learning disability. First, there is an attitude of discouragement about learning. The adolescent sees himself either as mentally retarded or as a slow learner. This identification in itself is built around the hidden logic, "I cannot learn to read (or write, spell, or do arithmetic) so why try?" Second, the psychopathology which has evolved from the long history of failure supersedes and masks the basic learning disability. The character structure which emerges is often far more difficult to modify than the learning disability itself. Third, because the adolescent has developed a way of life which incorporates a concept of being unable to learn, remedial efforts toward alleviating the deficit are frequently defeated (Anderson, 1970, p. 69).

It would appear logical that in order to help children cope successfully with their learning disability, we must begin before the nature of the disability has become obscured by emotional overlay. The message is clear for those children who

---

Lillian Whyte received her PhD from Columbia University, taught in elementary and secondary school in Nova Scotia and in the Department of Education at St. Francis Xavier University, Nova Scotia. She is currently with the University of Alberta as a Professor of Educational Psychology and Special Education.



are currently entering the school system at the kindergarten level. Nevertheless, there are thousands of adolescents who are learning disabled and who are now struggling to cope with mastering the requirements of the school system. For these young people we must continue to strive to find answers. The present study was an attempt to determine how learning disabilities are manifested during adolescence.

### *Characteristics of Learning Disabled Adolescents*

One of the conditions of a learning disability is that it persists past maturity (Kass, 1974) and is characterized by a persistence of chronicity (Rogan & Lukens, 1969). Longitudinal studies on severe reading disability indicate that the condition is best regarded as chronic, requiring long-term treatment (Balow & Blomquist, 1965; Silberberg & Silberberg, 1969).

Koppitz (1971), in a ten-year intensive investigation of 177 special-class students with severe learning disabilities, reported that at the end of five years only 30 children were able to return to regular class. The majority of the children, including those who returned to regular class, did not catch up in their work and their problems did not disappear. She concluded that there is an initial misleading spurt in achievement that arouses unwarranted optimism but that these early gains are often "washed out" in subsequent years.

Werner, Bierman and French (1971) reported on an impressive follow-up study of development and learning in which a cohort of more than 600 children and youth in Hawaii were followed from birth to 18 years; 88% of the original cohort were retained in the 18-year follow-up. Although of normal intelligence, approximately 3% of the original (1955) cohort were diagnosed in need of placement in a learning disability class at age 10 because of serious reading and communication learning disabilities, visual-motor impairment, hyperactivity, and difficulties in attention and concentration. Twenty percent of this group (the 3%) had physical evidence of "organicity" on pediatric-neurological examinations. For the overwhelming majority, serious problems persisted throughout adolescence. Agency records for 80% indicated continued academic underachievement confounded by absenteeism, truancy, and a high incidence of repetitive impulsive acting-out behavior. Rates of contact with community agencies were nine times as high for the experimental group as for the control subjects matched by age, sex, socioeconomic status, and ethnicity.

Group tests at age 18 showed continued perceptual-motor problems for most, as well as deficits in verbal skills and serious underachievement in reading and writing. Self-reports revealed a pervasive lack of self-assurance and interpersonal competency and a general inadequacy in utilization of intellectual resources on the Locus of Control Scale; the scores indicated the subjects did not feel their actions were under their own control. Other important highlights included:

1. Learning disabilities which persisted into young adulthood had strong biological and temperamental origins.
2. Although low SES increased the likelihood of learning disabilities, the combination of biological stress and environmental deprivations, especially family instability, led to greatest risk of subsequent serious and persistent learning disabilities in both lower and middle class children.
3. The quality of parent-child interaction played a pervasive role in the origins of learning disabilities, especially among unwed teenage mothers who are not

prepared physically, emotionally, or fiscally for adequately rearing their offspring.

4. The role of the father was quite crucial as a learning disabilities factor, especially as related to learning disabilities among sons and teenage pregnancies among daughters.
5. Major factors which contributed to the improvement of some learning disabilities included the youth's sense of competence and perceived internal locus of control as well as the ability to read and write standard English (Werner et al., 1971).

The question of how learning disabilities vary as a function of chronological age is a critical one because it has not been fully determined if characteristics commonly associated with learning disabilities at younger ages exist in adolescent populations. Rogan & Lukens (1969) stated that the disability may manifest itself in changing ways as the pupil matures and the school demands change. Clements (1966) listed, in order of prevalence, 10 differentiating characteristics of minimal brain dysfunctions in elementary school children: (1) hyperactivity, (2) perceptual-motor impairments, (3) emotional lability, (4) general coordination deficits, (5) disorders of attention, (6) impulsivity, (7) disorders of memory and thinking, (8) specific learning disabilities, (9) disorders of speech and hearing, and (10) equivocal neurological signs and encephalographic irregularities.

The literature indicates that learning disabled adolescents do not have a suppressed profile in all areas of functioning. Brutton, Richardson, and Magel (1973) stated that logical-reasoning and abstract-thinking skills do not develop as early in learning disabled pupils as they do in non-disabled students. Habertape (1976) noted that learning disabled students exhibit perseveration, disinhibition, and qualitatively different approaches to problem solving, while Deshler (1978) reported learning disabled adolescents may be deficient in the ability to organize information, deal with abstractions, questions, and evaluation of information. Stewart, Palkes, Miller, Young, and Welner (1974) reported several studies indicating that learning disabled children have lower than average IQs.

Hagin (1971) stated that the pervasive difficulty of the learning disabled student is an inadequacy in dealing with symbols which persists into adolescence. Hagin also concluded that learning disabled adolescents have difficulty with listening which interferes with social interactions, peer acceptance, job performance, and academic success. Deshler (1974) reported that these students performed poorly in blending tasks, spelling, grammar, and vocabulary skills. Myklebust (1973) found reading-disabled groups at 9, 11, 13, and 15 years were markedly deficient in output of written language, use of abstractions, the number of words available to them, and syntax, although the latter deficiency tended to lessen with age.

While perceptual and motor disabilities tend to improve with age, disabilities which still persist as significant during adolescence are fine and gross motor coordination problems (Siegel, 1974; Wilcox, 1970); handwriting difficulties involving illegibility, poor spacing, cramping, and laboriousness in writing (Deshler, 1974); clumsiness in physical education and disturbances in locomotor coordination (Myklebust, 1973); and perceptual irregularities.

Perceptual-motor correlates of learning disability also include hyperactivity, distractibility, poor attention, and lack of frustration tolerance. While the research is confusing and controversial, hyperactivity does not seem to disappear in



adolescence—in comparison to his peers, the adolescent remains more restless, distractible, impulsive, and emotionally unstable (Ross & Ross, 1976). According to Stewart (1976), it is possible that in adolescence, antisocial behavior appears for the first time in about one-fifth of hyperactive youth. The manifestations of hyperactivity, however, may change with age with a disappearance of “frantic-to-and-fro” purposeless motor activity and the appearance of subtler actions such as finger tapping, grimacing, or tics (Wilcox, 1970). Adolescent hyperactivity is of two types—distraction and attention lapses, or a paralysis of effort resulting in failure to concentrate (Wilcox, 1970).

Gordon (1970), Griffin (1971), Rosenthal (1973), and Silver (1974) have stressed that concomitant emotional and personality difficulties often are associated with learning disabilities in adolescents. Significant problems in social adjustment, social perception, self-concept, and motivation are often found. Socioeconomic and family variables also influence school achievement and related emotional and personality adjustments. Mauser (1974) reported negative self-concept and low frustration tolerance in learning disabled adolescents.

One of the most revealing and pernicious indicators of learning disability conditions during adolescence is the school dropout rate (Cervantes, 1965). Repeated academic failures affect the adolescent's personality. Among the social factors cited as significant are an aversive personality that repels rather than attracts people (Siegel, 1974) and a poor social perception, including factors such as an inability to generalize from one situation to another, oversensitivity to the reactions of others, difficulty in accurately interpreting both verbal and nonverbal moods and communications, and difficulty in determining the impact of one's actions on others. Repeated academic failure also leads frequently to a damaged self-concept and the development of defenses (Drake & Cavanaugh, 1970).

### *The Study*

Research supports the hypothesis that adolescents with learning disabilities continue to manifest dysfunction characteristics and emotional maladjustment to a significant degree. Furthermore, the deficits occur in various combinations in different students. The situation is more serious for adolescents than for younger children. By the time children reach adolescence, classes are departmentalized and teachers may not have the opportunity to observe the characteristics of their many students. The problem is compounded because many students have become very adept at covering up problems by clowning around, withdrawing, and generally using failure-avoiding tactics. Even with the availability of resources required to assess adolescents with test batteries designed to detect learning disabilities, because the tests are generally not suited to, or normed on, adolescent populations, the tasks of assessment and educational programming are gigantic.

One large vocational high school in Alberta, designed for adolescents having academic and social adjustment difficulties, includes students who were diagnosed as learning disabled in the elementary school. An effort was made to assess the learning disabled students for instruction more suited to their needs.

### *Subjects*

Three hundred thirty-one first year students in the vocational high school and 63 achieving grade 10 and 11 students in a regular high school were included. A review was made of the cumulative school records for each student and subjects

with IQ scores of more than 85 were selected. Because the students came from schools throughout the city and neighboring counties to the vocational high school, the IQ tests administered ranged from a number of group tests to the WISC-R. Therefore, no attempt was made to compute the mean IQ. All students were in the 15- to 17-year chronological age range.

### *Procedure*

Students completed a self-report inventory (Whyte, Harder & Kirk, 1981). The inventory was devised by observing the students in the vocational high school perform tasks in both the vocational and academic classes and by reviewing the literature on social adjustment characteristics usually manifested by learning disabled adolescents. From these sources a list of questions was compiled and revised both in content and phrasing in a series of pilot studies.

The revised inventory was then administered in groups of 15 to 20, to students from both schools. The questions in the inventory were grouped in six major categories (Table 1). Students were directed to read the questions and respond on a scale of 1 to 3: (1) they seldom or never experienced the problem; (2) they sometimes experienced the problem, and (3) they almost always experienced the problem. A category (4) was also included where appropriate to designate that the activity had never been tried. Students were encouraged to ask questions if they could not read and/or understand the question. Two or three examiners were always present. The inventory took approximately 40 minutes to administer.

### *Statistical Analysis*

Reliability for the inventory was determined by the test-retest procedure for a random sample of 60 subjects one month after the original administration. Items with correlation coefficients of 0.80 or more were retained for analysis. Responses for each item for both the vocational and academic high school students were compared and the Chi Square statistic were computed. One hundred fifty items discriminated between groups and were retained for the final analysis.

The average scale score for each area was determined by summarizing the scores recorded for each question in that area. For example, six items were retained in the area dealing with fine motor skills. A new scale from 1 to 5 was utilized to quantify the data: (1) 75% to 100% of the items were scored 1; (2) 50% were scored 1 and 50% were scored 2; (3) 75 to 100% of the items were scored 2; (4) 50% of the items were scored 2 and 50% were scored 3; (5) 75 to 100% of the items were scored 3. The summary scores for both groups for each area assessed were compared and the Chi Square statistic computed.

### *Results*

The percent of dysfunction (D) and nondysfunction (ND) students at each point on the scale are presented in Table 2. Percentages are rounded off to the nearest whole number and the totals may not always total exactly 100%. Because categories 1 and 2 on the 5-point scale indicated students were reporting minimal or no problems in an area, the percentages were combined and considered equivalent to 1 on the original scale. Category 3 on the 5-point scale was interpreted to mean that students reported the problem present sometimes; these percentages are included in category II of Table 2. Categories 4 and 5 on the 5-point scale were combined because the percentages in category 5 were very low. The combined percentages are



included under III in the table and are interpreted to mean that the student is reporting a significant problem in the area.

TABLE 1  
AREAS OF DYSFUNCTION ASSOCIATED WITH LEARNING  
DISABILITIES IN ADOLESCENTS

---

|  |
|--|
| I. SPATIAL DEVELOPMENT                       |
| 1. balance and posture                       |
| 2. laterality                                |
| 3. spatial judgments                         |
| 4. fine motor skills                         |
| 5. ocular control                            |
| 6. copying skills (visual-motor integration) |
| 7. visual perception                         |
| II. LANGUAGE AND COMMUNICATION               |
| 1. reception                                 |
| 2. comprehension                             |
| 3. expression                                |
| 4. understanding of grammar and syntax       |
| 5. auditory discrimination of sounds         |
| 6. sound blending skills                     |
| III. MEMORY/SEQUENCING                       |
| 1. memory                                    |
| 2. sequencing                                |
| IV. ATTENTION-RELATED DISORDERS              |
| 1. perseveration                             |
| 2. distractibility/impulsivity               |
| 3. hyperactivity                             |
| 4. emotional reaction to frustration         |
| V. ADJUSTMENT                                |
| 1. self-concept                              |
| 2. perception of how others see them         |
| 3. attitude to school                        |
| VI. SENSORY FUNCTIONING                      |
| 1. hearing                                   |
| 2. vision                                    |

---

Seven areas of spatial development were assessed. The majority of students in both groups did not report problems with balance and posture (91% and 98% for the D and ND groups, respectively). The Chi Square values are reported in Table 3. The differences reported for balance and posture between the two groups are not significant. The differences between the two groups for laterality were significant ( $p<0.05$ ). While 79% of the D and 90% of the ND students reported no problems, 10% of the former and 2% of the latter students reported significant problems with laterality.

TABLE 2  
PERCENTAGES OF DYSFUNCTION AND NONDYSFUNCTION  
STUDENTS AT EACH POINT ON SCALE

|                                    | Scale <sup>a</sup> |         |      |         |      |         |
|------------------------------------|--------------------|---------|------|---------|------|---------|
|                                    | I                  |         | II   |         | III  |         |
|                                    | DYS.               | NONDYS. | DYS. | NONDYS. | DYS. | NONDYS. |
| I. Spatial Development:            |                    |         |      |         |      |         |
| 1. Balance                         | 91                 | 98      | 5    | 2       | 3    | 0       |
| 2. Laterality                      | 79                 | 90      | 11   | 8       | 10   | 2       |
| 3. Spatial Judgments               | 68                 | 91      | 13   | 3       | 20   | 6       |
| 4. Fine Motor Skills               | 53                 | 78      | 13   | 10      | 34   | 13      |
| 5. Ocular Control                  | 72                 | 87      | 12   | 5       | 16   | 8       |
| 6. Copying Skills                  | 72                 | 93      | 16   | 3       | 12   | 3       |
| 7. Visual Perception               | 81                 | 100     | 9    | 0       | 10   | 0       |
| II. Language and Communication:    |                    |         |      |         |      |         |
| 1. Reception                       | 71                 | 91      | 17   | 8       | 12   | 2       |
| 2. Comprehension                   | 74                 | 96      | 12   | 5       | 14   | 0       |
| 3. Expression                      | 55                 | 88      | 19   | 6       | 26   | 5       |
| 4. Grammar & Syntax                | 75                 | 90      | 15   | 5       | 10   | 5       |
| 5. Discrimination                  | 83                 | 95      | 8    | 5       | 9    | 0       |
| 6. Sound Blending                  | 55                 | 76      | 19   | 11      | 26   | 13      |
| III. Memory/Sequencing:            |                    |         |      |         |      |         |
| 1. Memory                          | 57                 | 90      | 21   | 2       | 22   | 8       |
| 2. Sequencing                      | 69                 | 90      | 14   | 5       | 18   | 5       |
| IV. Attention Disorders:           |                    |         |      |         |      |         |
| 1. Perseveration                   | 68                 | 87      | 20   | 8       | 13   | 5       |
| 2. Distractibility/<br>Impulsivity | 52                 | 75      | 20   | 10      | 28   | 16      |
| 3. Hyperactivity                   | 59                 | 65      | 15   | 8       | 26   | 27      |
| 4. Frustration                     | 63                 | 87      | 9    | 0       | 27   | 13      |
| V. Adjustment:                     |                    |         |      |         |      |         |
| 1. Self-concept                    | 67                 | 86      | 15   | 3       | 18   | 11      |
| 2. Perception of Others            | 79                 | 92      | 12   | 5       | 9    | 3       |
| 3. Attitude to School              | 89                 | 100     | 7    | 0       | 5    | 0       |
| VI. Sensory Functioning:           |                    |         |      |         |      |         |
| 1. Hearing                         | 78                 | 95      | 12   | 3       | 10   | 2       |
| 2. Vision                          | 76                 | 90      | 8    | 0       | 16   | 10      |

<sup>a</sup> Scale: I - Problem seldom or never experienced  
 II - Problem sometimes experienced  
 III - Problem almost always experienced.

The two groups differed in their reported ability to make spatial judgments. Twenty percent of the D and 6% of the ND students reported significant problems in this ability. The differences were even more marked for fine motor skills: 34% of the D and 13% of the ND students reported significant problems ( $p < 0.001$ ). Significant difficulties were reported by 16% of the D and 8% of the ND students for ocular control ( $p < 0.05$ ).

Problems with copying skills, that is, the ability to write, reproduce figures, etc., were reported by 12% and 3% of the D and ND students, respectively ( $p < 0.01$ ).



TABLE 3

CHI SQUARE COMPARISONS OF DYSFUNCTION AND NON-DYSFUNCTION  
STUDENTS ON SIX CATEGORIES OF BEHAVIORS  
ASSOCIATED WITH LEARNING DISABILITIES

| Behavioral Areas                   | $\chi^2$ | d.f. | p   |
|------------------------------------|----------|------|-----|
| I. Spatial Development:            |          |      |     |
| balance and posture                | 4.233    | 2    | --- |
| laterality                         | 5.927    | 2    | *   |
| spatial judgments                  | 13.768   | 2    | *** |
| fine motor skills                  | 13.969   | 2    | *** |
| ocular control                     | 6.826    | 2    | *   |
| copying skills                     | 12.954   | 2    | **  |
| visual perception                  | 14.004   | 2    | *** |
| II. Language and Communication:    |          |      |     |
| reception                          | 11.027   | 2    | **  |
| comprehension                      | 14.688   | 2    | **  |
| expression                         | 26.017   | 2    | *** |
| grammar and syntax                 | 7.027    | 2    | *   |
| auditory discrimination            | 7.080    | 2    | *   |
| sound blending                     | 10.668   | 2    | **  |
| III. Memory/Sequencing:            |          |      |     |
| memory                             | 26.220   | 2    | *** |
| sequencing                         | 12.639   | 2    | *** |
| IV. Attention Related Disorders:   |          |      |     |
| perseveration                      | 9.615    | 2    | **  |
| distractibility/impulsivity        | 11.322   | 2    | **  |
| hyperactivity                      | 2.305    | 2    | --- |
| emotional reaction and frustration | 14.529   | 2    | *** |
| V. Adjustment:                     |          |      |     |
| self concept                       | 9.290    | 2    | **  |
| perception of how others see them  | 5.996    | 2    | *   |
| attitude to school                 | 8.005    | 2    | *   |
| VI. Sensory Functioning            |          |      |     |
| hearing                            | 10.461   | 2    | **  |
| vision                             | 8.517    | 2    | *   |

\*  $p < .05$   
 \*\*  $p < .01$   
 \*\*\*  $p < .001$   
 --- not significant

Finally, no ND student reported problems with visual perception and 81% of the D students did not report problems in this area, but 9% of the latter group reported a moderate problem and 10%, a significant problem ( $p < 0.001$ ).

Chi Square values in the areas of receptive language, comprehension, and expressive language were very significant ( $p < 0.01$ ). The percentages of D students reporting problems in the three areas were 12%, 14%, and 26% compared to 2%, 0% and 5% for the ND students.

Differences between the two groups in ability to understand grammar and syntax, and to discriminate auditorially were both significant ( $p < 0.05$ ) with 10% and 9% of the D and 5% and 0% of ND groups reporting problems in the two areas, respectively. Interestingly, both groups reported sound blending as a significant problem quite frequently (26% of D and 13% of ND); however, only 55% of the D students compared to 76% of the ND students reported no problems in this skill and the differences were significant ( $p < 0.01$ ).

Differences reported in both memory and sequencing abilities were very significant ( $p < 0.01$ ). Twenty-one percent of D students reported moderate and 22% reported significant problems in memory, and 14% and 18% of this group also reported moderate and significant problems with sequencing. Ninety percent of the ND students reported memory and sequencing were not problems.

The disorders usually included in the literature on hyperactivity are perseveration, distractibility/impulsivity, hyperactivity, and extreme emotional reaction to frustration. Differences for three of these behaviors were found to be significant in the present study. Differences in hyperactivity per se were not found to be significant; however, this was because both groups reported significant problems in this area: 26% of D and 27% of ND students scored 3 on the questions describing hyperactivity and the percentages at the other two points on the scale were also very similar.

Twenty-eight percent of D and 16% of ND students also reported significant problems with distractibility/impulsivity; however, 75% of ND but only 52% of D students reported having no problems in this area. The differences were significant for both distractibility and emotional reaction to frustration ( $p < 0.01$ ).

As reported previously in the literature review, emotional overlay appears to be the major disability among adolescents with learning disabilities. Three areas of functioning are usually implicated: negative self-concept, their perception of how they are viewed by others, particularly those in authority, and their attitude to school. The differences as determined by self-report of the D and ND students for these three areas were all statistically significant ( $p < 0.05$ ); however, negative self-concept was the area most often reported and negative attitude to school was cited frequently.

Finally, personal clinical experience indicates that learning disabled students at all age levels seem to have a higher than average incidence of problems associated with sensory functioning. Twenty-two percent of the D students reported moderate to significant problems with hearing compared to 5% of the ND students ( $p < 0.01$ ). Twenty-four percent of the D students reported moderate to significant problems with vision compared to 10% of the ND students ( $p < 0.01$ ).

### *Discussion*

The review of literature on the manifestations of learning disabilities during adolescence indicated that many of the characteristics of learning disabilities in children persist as significant problems into adolescence and adulthood. The present study offers support for this hypothesis.

Spatial deficits with the exception of balance and posture continue to interfere with the adolescent's functioning. The areas of most serious concern appear to be visual perception, fine motor skills in general, including the ability to write and to make judgments involving space. If the adolescent's perceptions of his spatial disabilities are accurate, there are serious implications. The majority of adolescents



with learning disabilities either elect or are counselled to enter vocational education, an educational stream loaded with potential for failure for individuals with spatial deficits. Furthermore, little has been done to investigate alternative teaching strategies in vocational education for children with these types of deficits. The area clearly is one where further research is essential.

The areas of language and communication deemed most serious by these adolescents involved listening, understanding or processing of visual and auditory information, and the ability to express themselves both vocally and manually. Sound blending was also considered an area of significant difficulty by 26% of these adolescents and of moderate difficulty by another 19%.

A review of the reading scores for these students reveals that many are functioning below the literacy level in reading. Reading material provided in shop classes is written above a grade six level. Remediation of reading relies heavily on a phonetic approach, including sound blending. Obviously, teaching strategies must be evolved which will first of all take into consideration the needs of the individual student, and secondly devise strategies in keeping with their pattern of strengths and deficits.

Forty-three percent of the dysfunction students reported moderate to severe disabilities in memory, while 32% reported problems in sequencing. If there is any one skill which separates the exceptional child from the achieving child, it is memory. Kirkbride (1980) compared the memory ability of achieving and learning disabled students at four age levels between 6 and 13 years of age and reported that the recall performance of learning disabled students at the 10- to 11-year age level lagged two years behind that of the achievers, while at the 12- to 13-year age level, the learning disabled students lagged four years behind the achievers in recall ability. However, when learning disabled students were taught to use verbal cues and grouping strategies for memory, the differences disappeared (Kirkbride, 1980). It is obvious that memory is involved in everything we do. It is equally obvious that learning disabled adolescents must be taught strategies to develop their memories.

The terms associated with the so-called syndrome of hyperactivity are often used imprecisely and synonymously; however, it is becoming more apparent that hyperactivity is not a single disability. Furthermore, those behaviors commonly associated with an increased activity level per se, for example, an increase in skeletal movement when the student should be quiet, are apparently found in both learning disabled and non-learning disabled students. However, on questions involving descriptions of perseverative, distractible/impulsive behaviors, and an intense reaction to frustration, the two groups do appear to be differentiated. Achieving students appear not to experience, or at least to cope more effectively, with these behaviors than do students with learning problems.

The behaviors described under attention disorders are probably also closely related to social/emotional adjustment. In combination with the other adjustment problems associated with learning disabilities in adolescents, it is little wonder emotional overlay problems take precedence over all other disabilities.

What are the implications of these findings? At the very least they indicate counselling is as important for these students as is instruction. Failure-avoiding tendencies must be overcome and a realistic and positive self-concept established, if these students are to become well adjusted and achieving adults.

Finally, the findings on sensory functioning may be indicating that many students are experiencing difficulty with vision and hearing. The need for adequate

diagnosis and correction, where possible, is essential. One wonders how many of these problems are going undetected.

### Conclusions

In summary, the present study provides support for the hypothesis that learning disabilities do not disappear during adolescence. Indeed, they persist as significant problems, affecting academic achievement, vocational instruction, and social and emotional adjustment.

It also appears that self-reporting is a viable approach in diagnosing the specific nature of disabilities among adolescents. At the very least, continued research in this area should make it possible to identify students with learning disabilities in our school population and make the more expensive and time consuming psychological testing more effective and efficient.

### References

- Anderson, R. P. (1970). *Guidance and the exceptional student: Guidance monograph series V*. Boston: Houghton Mifflin Co.
- Balow, B., & Blomquist, M. (1965). Young adults 10-15 years after severe reading disability. *Elementary School Journal*, 66, 44-45.
- Banks, I., & Finlayson, D. (1973). *Success and failure in the secondary school: An interdisciplinary approach to school achievement*. London: Methuen & Co.
- Brutten, M., Richardson, S. O., & Magel, C. (1973). *Something's wrong with my child*. New York: Harcourt Brace Jovanovich.
- Cervantes, L. F. (1965). Family background, primary relationships and the high school dropout. *Journal of Marriage and the Family*, 27, 218-223.
- Clements, S. D. (1966). *Minimal brain dysfunction in children*. NINDS monograph No. 3, U.S. Public Health Publication No. 1415. Washington, D.C.: U.S. Government Printing Office.
- Cruickshank, W. M. (1975). The psychoeducational match. In W. M. Cruickshank & D. P. Hallahan (Eds.), *Perceptual and learning disabilities in children: Psychoeducational practices*. Syracuse: Syracuse University Press.
- Deshler, D. D. (1974). *Learning disability in the high school student as demonstrated in monitoring of self-generated and externally generated errors*. Unpublished doctoral dissertation, University of Arizona.
- Deshler, D. D. (1978). Psychoeducational aspects of learning disabled adolescents. In L. Mann, L. Goodman, & J. L. Wiederholt (Eds.), *Teaching the learning disabled adolescent*. Boston: Houghton Mifflin.
- Doehring, D. (1968). *Patterns of impairment in specific reading disability*. Bloomington: Indiana University Press.
- Drake, C., & Cavanaugh, J. J. A. (1970). Teaching the high school dyslexic. In L. E. Anderson (Ed.), *Helping the adolescent with the hidden handicap*. Belmont, CA: Fearon Publishers/Lear Siegler.
- Gordon, S. (1970). Reversing a negative self-image. In L. E. Anderson (Ed.), *Helping the adolescent with the hidden handicap*. Belmont, CA: Fearon Publishers/Lear Siegler.
- Griffin, M. (1971). How does he feel? In E. Schloss (Ed.), *The educator's enigma: The adolescent with learning disabilities*. San Rafael, CA: Academic Therapy Publications.
- Habertape, J. F. (1976). *Problem solving in the learning disabled high school student*. Unpublished doctoral dissertation, University of Arizona.



- Hagin, R. (1971). How do we find him? In E. Schloss (Ed.), *The educator's enigma: The adolescent with learning disabilities*. San Rafael, CA: Academic Therapy Publications.
- Kass, C. E. (1974). Construction of a model of learning disability with attention to system-theoretic approaches. Proceedings of the International Federation of Learning Disabilities. Speech presented at the first meeting of the International Federation of Learning Disabilities, Amsterdam, the Netherlands.
- Kirkbride, A. J. (1980). *The development of classification and memory ability in achieving and learning disabled children: Piagetian and levels of processing approaches*. Unpublished doctoral dissertation, University of Alberta.
- Koppitz, E. M. (1971). *Children with learning disabilities: A five year follow-up study*. New York: Grune & Stratton.
- Mauser, A. J. (1974). Learning disabilities and delinquent youth. *Academic Therapy*, 6, 389-402.
- Meier, J. H. (1976). *Developmental and learning disabilities: Evaluation, management and prevention in children*. Baltimore: University Park Press.
- Myklebust, H. R. (1973). *Development and disorders of written language: Studies of normal and exceptional children*. New York: Grune & Stratton.
- Rogan, L. L., & Lukens, J. E. (1969). Education, administration and classroom procedures. In N. G. Haring (Ed.), *Minimal brain dysfunction in children* (NINDS Monograph No. 2 U.S. Public Health Communication No. 2015). Washington, D.C.: U.S. Government Printing Office.
- Rosenthal, J. H. (1973). Self-esteem in dyslexic children. *Academic Therapy*, 9, 27-39.
- Ross, D., & Ross, S. (1976). *Hyperactivity: Research, theory, and action*. New York: Wiley.
- Siegel, E. (1974). *The exceptional child grows up*. New York: E. P. Dutton.
- Silberberg, N., & Silberberg, M. (1969). Myths in remedial education. *Journal of Learning Disabilities*, 2, 209-217.
- Silver, L. B. (1974). Emotional and social problems of children with developmental disabilities. In R. E. Weber (Ed.), *Handbook on learning disabilities*. Englewood Cliffs, NJ: Prentice-Hall.
- Stewart, M., Palkes, H., Miller, R., Young, C., & Welner, Z. (1974). Intellectual ability and school achievement of hyperactive children, their classmates and their siblings. In D. Ricks & M. Roff (Eds.), *Life history research in psychopathology* (Vol. 3). Minneapolis: University of Minnesota Press.
- Stewart, M. (1976). Is hyperactivity abnormal? And other unanswered questions. *School Review*, 84, 31-42.
- Werner, E. E., Bierman, J. M., & French, F. E. (1971). *The children of Kauai*. Honolulu: University of Hawaii Press.
- Whyte, L., Harder, E., & Kirk, P. (1981). Inventory for detection of learning disabilities in adolescents: A self-report. Edmonton, Alberta; Edmonton Public School Board. Unpublished report.
- Wilcox, E. (1970). Identifying characteristics of NH adolescents. In L. E. Anderson (Ed.), *Helping the adolescent with the hidden handicap*. Belmont, CA: Fearon Publishers/Lear Siegel.

MARVIN F. WIDEEN

*Simon Fraser University*

and

DAVID HOPKINS

*West Glamorgan Institute of Higher Education*

## Supervising Student Teachers: A Means of Professional Renewal?

*The professional growth occurring among teachers as a result of their association with the York Faculty of Education was assessed through a case study. Interviews and a questionnaire distributed to 325 teachers and principals in 31 schools were analysed to determine the occurrence of professional growth, the nature of that growth, and the factors contributing to it. The results support only in part the commonly held idea that association with a university teacher training program promotes professional growth. Only in those schools where some compatibility existed with the program at the university and where teachers felt involved in that program was evidence found for the occurrence of professional growth. The main source of influence was the student.*

The need for teachers and others in various social institutions to engage in professional growth to keep pace with social change is well documented (Dobson & Dobson, 1978; Flanders, 1980; Howsam, 1976; McLaughlin & Marsh, 1978). Where programs of inservice education for teachers have been analysed, however, the results have been less than encouraging. In a paper presented to an OECD conference on inservice, Fullan (1980) argued that much of the current inservice activity is ineffective because the school environment is ignored, evaluation and follow up seldom occur, and conceptual planning is lacking. Agreement exists that the one shot workshop, although widespread, is almost totally ineffective in terms of assisting the professional development of teachers.<sup>1</sup>

---

Dr. Marvin Wideen is an Associate Professor of Education at Simon Fraser University. He recently codirected a large-scale research project on the management of change in teacher education in Canada and is conducting a program of research into the professional development of teachers.

Dr. David Hopkins is a Lecturer in Educational Research at the West Glamorgan Institute of Higher Education in South Wales. He recently graduated from Simon Fraser University, where he codirected an institute on school improvement.



Initiatives designed to redress the weaknesses in current inservice programs for teachers tend to focus on school based programs that provide continuous activity over time. Henderson (1979) argues that a school focus concentrating on problems identified at the school level and drawing on outside expertise combines the best aspects of course-based and school-based inservice activity.<sup>2</sup> Continuity over time has long been espoused as an important feature of inservice programming. The recent work of Joyce and his associates (Joyce, 1979; Joyce & Showers, 1980) has provided the main conceptualization of this notion. They argue that five training components—theory presentation, modeling, practice, feedback, and coaching—are required for teachers to learn and effectively apply new knowledge.

The resources necessary for initiating school-based inservice programs and/or the training models proposed by Joyce are simply not to be found in most school or university jurisdictions. Consequently, implementation of these ideas on any meaningful scale is highly unlikely in the near future.

Little consideration in the literature on inservice seems to have been devoted to those structures which are in place and which have promise for professional growth among teachers. One area that has the potential to provide those elements of school-focused inservice identified by Henderson (1979) and to provide continuity over time is the cooperating teachers' association with faculties of education during their supervision of student teachers. It has been assumed both by those in the schools and those in Faculties of Education<sup>3</sup> that this association with students and university faculty will result in professional growth for teachers, but this claim remains largely unsupported. It would seem a useful contribution to the general literature on inservice, therefore, to study the association between a university and its cooperating teachers with the view to assessing the professional growth that occurs during the practicum and the conditions under which such growth can be maximized.

In this paper we assess the degree of professional growth occurring among teachers who worked with the Faculty of Education of York University supervising students during practice teaching. Three questions are addressed: (1) Is inservice activity and professional growth that can be attributed to the university seen to be occurring in the schools and among teachers? (2) What is the nature of that activity and growth? (3) What factors contribute to it?

Before turning to procedures and a reporting of results, we discuss related literature and outline the York teacher education program focusing on the roles of the adjunct professor and host teacher. In the final section, we discuss the results in terms of the questions and issues that direct the study.

### *Relationship to the Literature*

The literature on inservice education, with its focus on proposals and arguments for new and more effective ways of retraining teachers, typically ignores the implementation of these programs. Consequently, such initiatives are likely to have little impact on what goes on in classrooms. We have argued that an alternate and potentially more useful route is to examine and seek to improve existing areas that provide opportunities for professional growth for teachers.

A small cluster of studies have focused on the role played by cooperating teachers in the supervision of student teachers during the practicum. In a recent review of research into teacher education in Canada, Holborn and Wideen (1983) identified 14 reports dealing with the cooperating teacher, of which several have

implications for this study.<sup>4</sup> One consistent finding from these studies is that the role of the cooperating teacher is poorly defined and that, generally, teachers are unprepared for the task of supervision (Caldwell, 1979; Ratsoy, Babcock & Caldwell, 1978; Wideen, 1973; Holborn, 1978). Not surprisingly, then, teachers' reactions to various approaches used in assisting them to improve supervisory skills has been highly positive (Sloan, 1981; Holborn, 1978; Richards & Stinson, 1975).

The extent to which teachers learn new skills and gain new ideas from the university and are also able to implement them in classrooms remains something of an open question. For example, Marble (1981) summarized several studies conducted at Simon Fraser and reported that only parts of the model of clinical supervision used at that institution were being applied by the cooperating teachers. This was not the case, however, in a study carried out at Edmonton by Perfrement and Sigurdson (1978) who reported that teachers simply did not use clinical supervision even with extensive training. On the other hand, Martin, Isherwood and Rapagna (1978) found that regardless of the approach, training in supervision had a positive effect on teachers and the students they supervised.

What appears missing in these studies is an attempt to identify what professional growth occurs among teachers as a result of the programs sponsored by the university and equally important, what conditions are related to that growth. Another area not addressed in any of the studies we reviewed is the sources of influences upon teachers which appear most potent in terms of influencing their behavior. For example, it is not at all clear why Perfrement and Sigurdson report that the teachers in their study simply did not use clinical supervision while Martin and others found that training in supervision regardless of the approach was effective. While the activities sponsored by York University are not exactly the same as those discussed in this review, the results from our study will shed light on the same questions.

Indirectly the study should contribute to the growing literature on the problems of integrating theory and practice. Evidence exists to support the notion that this problem is one of the most serious in teacher education. The problem is exemplified by the different perspectives taken by those in the university and school communities as to each other's roles in preparing teachers. Those in the school may see faculties of education as being out of touch with the reality of public education, bound by intellectual theorizing and blissfully oblivious to the practical difficulties of classroom teaching. The university staff, on the other hand, may regard with some disdain what they consider the folklore and common knowledge of the classroom; the view is taken that instruction would be greatly improved if only what is known about teaching was actually implemented (Allen, 1976). It is the effective blend of these two cultures or stereotypes that should be the aim of teacher training programs.

### *The Setting*

The establishment of York University in the 1950s and the planning for a faculty of education in the mid-sixties presented an opportunity for a new approach to teacher education in Ontario. Following a lengthy development period involving a thorough discussion and analysis of present teacher education, a model was implemented in 1972. The prime objective was to produce graduate teachers who were capable "both of working in emerging schools and of effecting change in schools" (Overing, 1973, p. 19). To this end, the faculty planners sought to redefine



their relationship with the schools; to enter with them into an active partnership in the profession. Under this scheme, it was hoped that inservice and preservice training would become natural extensions of each other and that the role of both cooperating and student teachers would radically alter. It would not be an exaggeration to say that the York program was undertaken with a definite sense of mission.

This sense of mission pervaded the teacher education program at York which was distinguished by four dominant characteristics. First, teacher education was seen to be a responsibility of the whole university where the majority of education faculty were cross appointments from within the university. Second, students were to co-register in two degree programs so that they pursued an academic degree and the professional B.Ed. simultaneously over a five-year period. Third, the curriculum of the B.Ed. program made extensive use of practicums, field experiences, and related seminars. The students spent one day a week in a school during the academic year culminating in a three- to five-week practicum; throughout this time they were associated with one particular host teacher. Fourth, the practicums and seminars involved the participation of selected school teachers who served as adjunct professors seconded to the university. These adjunct professors acted both as seminar leaders and supervisors of the student teachers' professional development.

In the development of their program, the planners at York saw the emergence of two roles vital to their conception of teacher education, namely that of the adjunct professor and the host teacher. It is these roles which provide the focus of the study. The *adjunct professors*, envisaged as a bridge between the university and the school, were successful teachers seconded to York for one-half day each week. In the school, they worked with the host teachers and carried out a weekly evaluation of students. They also engaged in a certain amount of professional development conducted in the school among the host teachers and students, thereby bringing the content and the theoretical input from the university to the school setting. On the York campus, adjunct professors worked with a faculty member in organizing and conducting seminars for students. The specific content of those seminars was based on joint input from the adjunct professor and the faculty person. The topic selection was also driven by the needs of students and issues raised by them.

One intention of the planners at York was to cooperate with schools that had a history of innovation and were amenable to change; therefore, the role of the *host teacher* was critical. Student teachers, called educational assistants, were assigned to a selected host teacher to serve as an apprentice for one day a week throughout the academic year for the full duration of their degree program. The expectation was that the student teachers would enter into a long-term association with the school and the host teacher, and that they would devise and implement an instructional program with a diagnostic, prescriptive, and evaluative emphasis. By emphasizing the collegial interaction of the student teacher and host teacher working together in a school that was attempting to modify its instructional program, the elements of inservice and preservice training could be merged. It was hoped that, while host teachers might lack sufficient skills to enable them to implement their ideas, interaction with the York program would aid them to acquire these skills.

It was within this context that this study, which focussed on the roles of the adjunct professor and host teacher, was conducted. While these roles were defined

in terms of the York program and as such are unique, the roles possess certain similarities to those of cooperating teachers in other faculties of education in Canada. The role of adjunct professor is not unlike that played by teachers seconded to universities for periods of one or two years. Host teachers as well do many of the things cooperating teachers do in supervising students during an extended practicum which is now a common practice in many Canadian universities.

### *Procedure*

A case study approach was used in seeking answers to the questions identified in the introduction. The principal investigator spent two one-week periods at York and a research assistant conducted interviews over a similar period. Data were obtained through structured interviews and a questionnaire administered to principals, host teachers, and adjunct professors. Throughout the study, the process of triangulation was used to confirm perceptions emerging from different sources of data and different methodological techniques.

To determine whether professional growth among teachers and inservice activity in schools were occurring and to learn something of their nature (the first two questions raised by the study), a questionnaire was used, followed up by an interview. On the questionnaire, the respondents (host teacher, principals, and adjunct professors) were asked to note on a four-part scale the extent to which the operation of their classrooms and their methods of teaching had been influenced by the York Teacher Education program. In addition, an open-ended question asked the same groups to describe things they did differently as a result of their association with York. These data were followed up by interviews and classroom visits. Tape recordings of interviews were analysed to confirm the data collected from the questionnaire. Wherever possible, cross-checks were made by asking other teachers and principals to confirm information. The data from these sources were then matched against five criteria which are indicative of the presence of inservice activity and professional development. These criteria, outlined below, were drawn from the literature and from the general experience of those working in the area of inservice and were set out prior to the first field visit.

1. *Transactional activity.* Evidence should exist that traditional inservice activities such as workshops, reading and after-school programs are occurring.
2. *Perception of growth.* There should be a perception of growth and change in behaviour both among the individuals concerned and those who associate with them.
3. *Instructional and organizational change.* Evidence should exist that instructional and organizational changes in the classroom or workplace have in fact occurred.
4. *Less isolation among teachers.* There should be evidence that teachers were engaging in collegial exchange and planning and that the tendency to work in isolation was breaking down.
5. *Self-criticism.* Teachers should show evidence that they were becoming more critical of their own work.

To address the third question of the study, those factors related to professional growth and inservice had to be identified. It was assumed at the outset of the study that variation in terms of both inservice and professional growth would occur



among teachers associated with York. Initial contact with the schools and selected interviews confirmed this variability. Thus the questionnaire and interviews were structured to identify differences among them. These differences we discuss as factors related to professional growth and inservice.

Results

1. Does the York Teacher Training Program lead to inservice education activity and to professional development?

Cooperating teachers (adjunct professors and host teachers) were asked the extent to which their teaching (classroom operation and methods of teaching) had been influenced by their association with York. Principals responded to the same question but from the perspective of the teachers in their schools. The responses to this question, which are provided in Table 1, gave an indication of whether or not teachers perceived themselves as influenced by the York program. Considering only the aggregate data in Table 1, it can be seen that 56% of those responding reported a great deal or a moderate amount of influence and that only 13% reported not having been influenced. It can also be seen that differences in perception existed between the adjunct professor, host teacher, and principal; for example, 73% of the adjunct professors reported having been influenced moderately or a great deal, while only 51% of the host teachers responded in these two categories.

TABLE 1

RESULTS TO QUESTION B1 ASKING THE EXTENT OF INFLUENCE OF THE YORK TEACHER EDUCATION PROGRAM ON CLASSROOM OPERATION AND TEACHING METHODS<sup>a</sup>

| Subgroup          | Response Scale |                   |             |            | Total |     |
|-------------------|----------------|-------------------|-------------|------------|-------|-----|
|                   | A Great Deal   | A Moderate Amount | Very Little | Not At All | %     | N   |
| Principal         | 12%            | 58%               | 24%         | 6%         | 9%    | 17  |
| Adjunct Professor | 8%             | 65%               | 28%         | 0          | 17%   | 29  |
| Host Teacher      | 8%             | 43%               | 33%         | 13%        | 74%   | 128 |
| Total % (N)       | 8% (14)        | 48% (84)          | 31% (54)    | 13% (22)   | 100%  | 174 |

<sup>a</sup>The data in this, and other tables are based on a 56% return of the questionnaire administered to 325 teachers and principals in 31 schools.

In a question soliciting written comments, teachers were asked if they could describe anything specific that they did differently in their classrooms as a result of association with York. These comments were categorized and are reported in the first column in Figure 1. Those teachers and principals who were interviewed were asked if there had been benefits in associating with York and, if so, what they had been. The responses to this question, which were categorized from transcribed interviews, are reported in Figure 1, column II. When the results from these two categories are considered with the criteria cited previously, a number of relationships become evident, as can be seen in the figure.

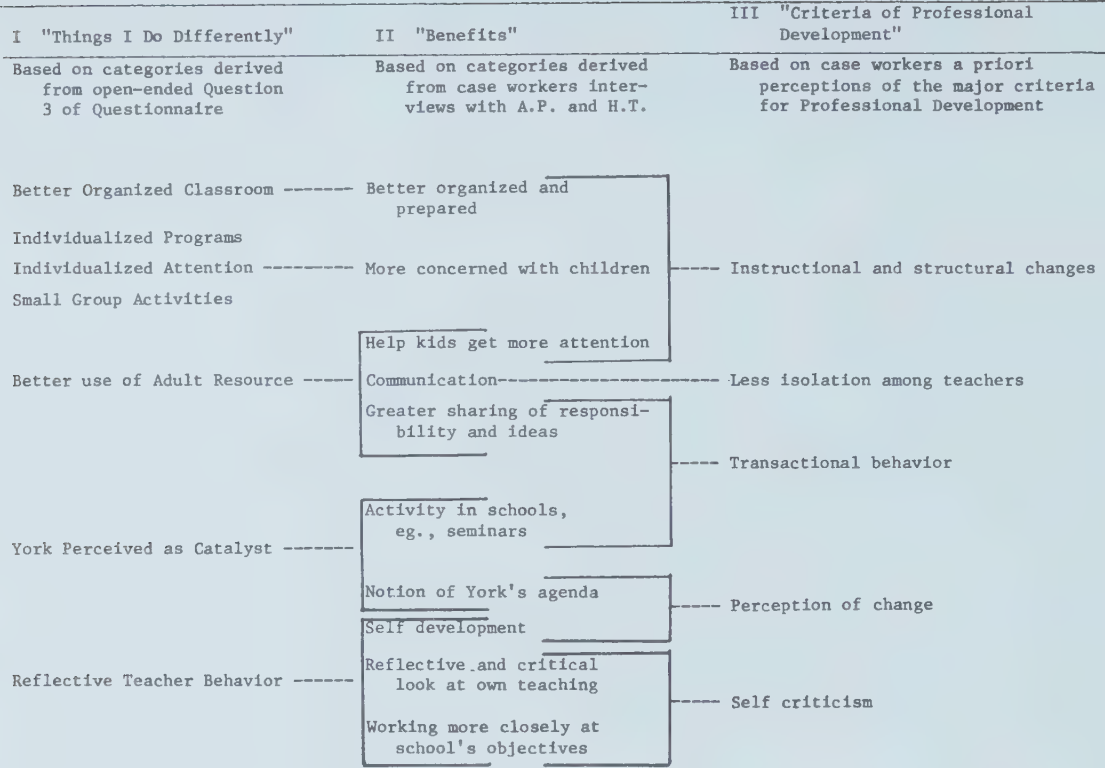


Figure 1. Relationships between teacher's perceptions and the criteria for professional development.

For example, the three categories of open ended responses, individualized programs, individualized instruction, and small group activity, which fell under *things I do differently* in column I, all involved individual differences. As a group they were related to the category, *more concerned with children*, under *benefits* in column II. These matched the professional development criterion of instructional and structural change in column III. Thus, the three columns in Figure 1 demonstrate the interrelationships between the responses that emerged from the different data and their relationship to the criteria. It is the relationship between the categories in columns I and II and the criteria listed in column III, coupled with the results described earlier in Table 1, that is the basis for the general claim that inservice activity and professional development which could be linked to York did in fact occur in the perception of teachers and principals.

2. What was the nature of the professional development?

To obtain an indication of the nature of professional development, host teachers and adjunct professors were asked their level of agreement on eight types of influence which were identified during the pilot interviews. The percentage results of these eight items for the total group, adjunct professor and host teacher are provided in Table 2.

It can be seen that considerable variation occurred for the different sources of influence. Two categories in particular, observing individual differences and responding to individual differences, carried a noticeably high level of agreement. This finding is significant insofar as it matches the York faculty's emphasis on the individual needs of children. This match is particularly interesting in the light of other information (discussed later) which emerged during interviews where teachers seemed unaware of any particular philosophical emphasis on the part of York. The



high level of disagreement for the item, *staff relations changing*, was also of significance considering that the faculty at York considered this a likely outcome of their model. As noted in the discussion of the results shown in Table 1, differences in perception between the host teachers are in evidence; for example, for the first item *modifying instructional goals*, 28% of the adjunct professors agreed that they had been influenced while only 13% of the host teachers expressed that view. On all but one of the items, the adjunct professors indicated a higher level of agreement. The exception was in responding to individual differences.

TABLE 2  
LEVEL OF AGREEMENT GIVEN BY ADJUNCT PROFESSORS  
AND HOST TEACHERS TO DIFFERENT SOURCES OF INFLUENCE

| Types of Influence                   | Response Scale |      |       |         |    |       |          |    |       | Number Responding |    |     |
|--------------------------------------|----------------|------|-------|---------|----|-------|----------|----|-------|-------------------|----|-----|
|                                      | Agree          |      |       | Neutral |    |       | Disagree |    |       | Total             | AP | HT  |
|                                      | AP             | HT   | Total | AP      | HT | Total | AP       | HT | Total |                   |    |     |
| Modifying Instructional Goals        | 28*            | 13** | 17*** | 34      | 28 | 29    | 38       | 59 | 54    | 127               | 29 | 98  |
| Modifying Teaching Methods           | 32             | 25   | 27    | 36      | 25 | 27    | 37       | 50 | 46    | 128               | 28 | 100 |
| Observing Individual Differences     | 86             | 66   | 71    | 4       | 10 | 9     | 11       | 24 | 20    | 126               | 28 | 98  |
| Responding to Individual Differences | 52             | 59   | 57    | 32      | 13 | 17    | 16       | 28 | 26    | 129               | 25 | 104 |
| Changing of Staff Relations          | 32             | 19   | 15    | 22      | 15 | 17    | 46       | 25 | 26    | 125               | 28 | 97  |
| Work More Challenging                | 42             | 36   | 36    | 26      | 28 | 27    | 32       | 36 | 37    | 128               | 29 | 99  |
| Work More Satisfying                 | 33             | 43   | 41    | 48      | 38 | 34    | 19       | 27 | 25    | 120               | 27 | 103 |
| Make Better Use of Resources         | 58             | 41   | 44    | 21      | 28 | 28    | 21       | 33 | 31    | 131               | 29 | 103 |

Note. HT = Host Teacher AP = Adjunct Professor

\* 28% of the 29 APs responding  
\*\* 13% of the 98 HTs responding  
\*\*\* 17% of the total number of 127 responding

3. What factors maximize inservice activity and professional development?

As indicated earlier it was apparent from the outset that some teachers and some schools had been influenced by their association with York and others had not. Our initial perceptions were confirmed by the first item on the questionnaire which asked specifically about the influence of York. The responses in each school were aggregated for this item, the schools ranked on this basis, and the data reexamined in order to identify factors related to the difference in influence. From this analysis, three factors emerged which appeared to maximize the effects of York’s programs—involvement, compatibility, and the influence of students.

Involvement, defined as participation in York’s activities, was a factor both for individual teachers and for schools. Those who spoke of having been influenced by

York's program or who could provide examples which exemplified that influence were those who had a record of participation in activity associated with York. In the case of individual teachers, the distinction was seen most clearly when the roles of the adjunct professor and the host teacher were contrasted. As earlier described, the adjunct professor's role carried with it much greater involvement than that of the host teacher. However, as shown in Table 1, the adjunct professors reported having been influenced to a greater extent and expressed a higher level of agreement on six of the eight types of influences shown in Table 2.

The second factor related to the occurrence of inservice activity and professional growth was that of compatibility. As we inquired into the notion of compatibility it became evident that roughly three groups of schools existed with regards to the staffs' perceptions of York's concept of teacher education—those school staffs who were in agreement with York's mission, those who were not, and those who saw York as just another Ontario teacher's college. It was quite evident from our interviews that it was within the first group—those schools in which staffs were attuned to York's mission—that the most inservice activity and professional growth were occurring. Much less professional growth or inservice activity could be identified in the other two groups of schools. While these two factors, compatibility and involvement, were related, it was clearly evident that they were separate particularly in the two groups of schools whose staffs were either attuned to York's mission or opposed to it.

Another factor highly related to inservice activity and professional development was the source itself. In early interviews it became evident that York's association with the schools came about through the complex interaction of roles played by different actors. Each represented a potentially different source of influence. Six sources (York students, peer interaction, principal, self-examination, the York staff, and the adjunct professor) were identified on the questionnaire; both the adjunct professor and the host teacher were asked to select a level of importance for each. Eighty per cent of the adjunct professors and 68% of the host teachers identified the presence of York students in their classroom as being the most important source of York's influence upon them. The next most important influence was self-examination reported by 56% and 63% of the adjunct professors and host teachers, respectively. The principal and the York staff were identified as the least influential of the six sources by the adjunct professor and host teacher respectively.

### *Discussion*

A common notion about student teaching held by those in faculties of education and in school districts is that teachers who supervise students grow professionally. The data from this study do not support that commonly held belief. What they do support is the notion that teachers who are involved in a faculty's program and whose view of education is compatible with that of the faculty do grow professionally. To others, the task of student supervision is just that—a task to be carried out. If professional development is to be a by-product of student supervision, then both schools and faculties of education, on that basis, must provide opportunities for teachers to become fully involved as active partners based on a mutual understanding of goals.

The main influence perceived by teachers to have resulted from their association with York involved observing and responding to individual differences. It is noteworthy that prior to and during the time of the study, the faculty at York



perceived the recognition of individual differences as being an important component of effective teaching. Such data suggest, then, that not only can faculties of education encourage professional growth, but that they can influence the direction of that growth. However, influence upon teachers is a complex process and not an automatic one. We saw several schools in which virtually no inservice activity attributable to York could be identified. One of the factors differentiating the two groups of schools was the compatibility between the schools' orientation and that of York.

The notion of compatibility as a factor related to effective influence raises questions about the practice in most faculties of education whereby students are confronted with a bewildering array of alternatives from which they are to somehow make up their minds as to what constitutes effective practice. York, on the other hand, espoused a distinct idea of practice which was reflected in the selection of co-operating schools, the focus of seminars, the provision of materials, and the legitimization of practice. Consequently, students entered the schools with a fairly clear perception of what (in York's terms) constituted good education. In those schools where the staff's view of education was compatible with that at York, the students found a receptive climate; it was probably no coincidence that it was in those schools that most inservice was seen to be occurring and from which most professional development was reported. As a corollary, it is questionable whether the stance taken by most other faculties of education enables them to be influential in producing professional growth.

The finding that students represent a very strong influence upon teachers was quite clearly demonstrated in this study and has been supported by data collected at Simon Fraser University (Terry, 1980) where 65% of a sample of teachers, when asked how association with that university had influenced their professional development, identified student teachers as the main source of influence. Apparently, the influence of the students caused self-analysis as well as a clarification and stimulation of ideas. The finding of a strong student influence is not surprising in itself. The fact that the influence of students was seen to be greater than that of the principal or the faculty members is surprising and perhaps somewhat disconcerting.

One might attribute such influence to the constant presence of students in the classroom. However, the lack of threat posed by a student as distinct from a principal or university person should also be considered as an explanation of the finding. One might also speculate that such a finding simply underlines the isolation, loneliness, and intensity of the teaching workplace. Other potential sources of influence such as the principal, a peer, or a university supervisor simply do not penetrate the day to day activity in which a teacher must engage. The student, on the other hand, does.

Certainly, the two cultures to which we referred earlier with respect to theory and practice were alive and well at York, at least on the surface. But there was ample evidence to suggest that just as the practice of the school was being modified by the theory of the university, the reverse was also true, and that York's ideas were also being modified and adapted. This led us to the hypothesis that a hidden curriculum was operating at York and was exercising a substantial influence on many of the schools. It was working through a vector of forces—the most powerful of which was the student—to cause teachers to look more closely at what they were doing and to implement certain practices such as individualized instruction. As

such, the problem of theory and practice was being addressed, albeit in a rather indirect and subtle way. Thus the data from this study suggest that the type of school-based teacher education program typified by York has been somewhat more effective in providing an adequate blend of theory and practice than have other programs.

The conclusion to which one is drawn by the results of this study is that the model of teacher education operating at York provides impetus for professional development among many of the teachers associated with it. For others, it clearly provides very little except the added burden of having student teachers. The conditions had to be right: teachers had to be involved and have a stake in the program; the theoretical message had to be consistent and the atmosphere of the school receptive to that message; and the sources of influence had to be close to the action. Contrast this ideal scenario with a typical student teacher practicum where schools are often picked out of expediency, where no activity is planned to aid the supervising teacher, where meaningful contact with the university is minimal, and where the students leave the university confused by a plethora of ideas, none of which have been explored in depth. The message is clear. If teacher training institutions are to provide professional growth for those in the schools upon whom they depend to supervise their students, a great deal more thought and planning must go into the working relationship between the typical university faculty of education and the school.

#### Notes

1. The reader is referred to the work of Joyce and Showers (1980) and Joyce (1979), a review by Young (1979), and a review by Wideen (1980).
2. Henderson's position has been echoed and further developed by several writers. See in particular, Fullan (1980) and Nur and Weil (1978).
3. The authors have analysed the comments of school district persons and faculty members who discuss student teaching. A theme in these discussions, and indeed a rationale for district cooperation, is the notion that teachers will benefit professionally from the experience of supervising a student.
4. This review (Holborn & Wideen, 1983) is one of a four-volume report on a study of teacher education conducted in Anglophone Canada. Copies of the full report are available at a cost of \$10.00 from M. Wideen, Simon Fraser University, Faculty of Education, Burnaby, B.C. V5A 1S6.

#### References

- Allen, I. (1976). Extending the practicum: Problems in integrating theory and practice. *Canadian Journal of Education*, 1(3), 43-51.
- Bolam, R. (1980). In service education and training. In E. Hoyle (Ed.), *World yearbook in education 1980: Professional development of teachers*. New York: Nichols.
- Caldwell, B. J. (1979). *The evaluation of an education practicum: Implications for field personnel*. Paper presented at the annual meeting of the Canadian Society for Study in Education, Saskatoon.
- Dobson, R., & Dobson, J. (1978, August). *The university and in service education in public schools: Instructional strategies and delivery system*. Paper presented at Deans Network Conference, Purdue.
- Flanders, T. (1980). *The professional development of teachers*. Commissioned report. Vancouver, B.C.: British Columbia Teachers Federation.
- Fullan, M. (1980). *School focused in-service education in Canada*. A report prepared for OECD, Paris: OECD.



- Henderson, E. S. (1979). The concept of school-based inservice education and training. *British Journal of Teacher Education*, 5(1), 17-25.
- Holborn, P. (1978). *The role of the cooperating teacher*. Paper presented at the annual meeting of the Westcast, Winnipeg.
- Holborn, P., & Wideen, M. (1983). *Research in Canadian teacher education: A review (Volume IV)*. Project report. Burnaby: Simon Fraser University.
- Howsam, R. (1976). The profession of teaching. In *Issues in inservice education*. Syracuse, NY: National Council of States on Inservice Education.
- Joyce, B. (1979, May). In-service: New perspectives on an old term. In M. Wideen, D. Hopkins & I. Pye (Eds.), *In service: A means of progress in tough times*. Conference proceedings. Burnaby: Simon Fraser University.
- Joyce, B., & Showers, B. (1980, February). Improving inservice training: The message of research. *Educational Leadership*, 375-385.
- Marble, W. O. (1981). *Evaluation in the professional development program at Simon Fraser University*. Paper presented at the annual meeting of the North West Association of Teacher Educators, Burnaby.
- Martin, Y. M., Isherwood, G. B., & Rapagna, S. (1978). Supervisory effectiveness. *Education Administrative Quarterly*, 14(3), 74-88.
- McLaughlin, M., & Marsh, D. (1978). Staff development and school change. *Teacher's College Record*, 80(3), 69-84.
- Nur, M., & Weil, M. (1978). Staff development models: A report on the first year. Report for California State Department of Education, ESEA Title IVC.
- Overing, R. (1973). Toward a redefinition of teacher education. *Interchange*, 4(2-3), 19-27.
- Perfrement, E., & Sigurdson, S. (1978). An evaluation of cooperating teacher workshops—Ed. Pra. 252—and a course evaluation of Ed. Pra. 252. In Ratsoy, E. W., *Program evaluation activities of the Faculty of Education, The University of Alberta*. Program Evaluation Report No. 5. Edmonton: Faculty of Education, Office of Program Evaluation.
- Ratsoy, E., Babcock, G. R., & Caldwell, B. J. (1978). *Evaluation of the education practicum program 1977-1978*. Program Evaluation Report No. 1. Edmonton: Faculty of Education, Office of Program Evaluation.
- Richards, L., & Stinson, E. W. (1975). A preliminary project in intern supervision in selected Saskatchewan high schools. *Saskatchewan Journal of Educational Research and Development*, 5(2), 38-45.
- Sloan, L. V. (1981). *Final report on the development of the extended practicum at the University of Alberta*. Edmonton: University of Alberta, Office of Program Evaluation.
- Terry, B. (1980). *Summary of interviews with school associates*. Project report, Burnaby: Simon Fraser University.
- Wideen, M. F. (1973). *A report on an evaluation of education 402 conducted in the fall semester. No. 1*. Unpublished report. Burnaby: Simon Fraser University.
- Wideen, M. F. (1980). *In service education in the district of North Vancouver: A case study*. Unpublished manuscript, Simon Fraser University.
- Young, E. (1979, July). *Inservice education: A guide to effective practice*. Unpublished report. Regina: Saskatchewan Teachers' Federation.

MARILYN WESTBURY

*The University of Alberta*

## The Link Between Universities and Employment in Alberta

*The study has four major thrusts: first, it presents university supply/labor market demand imbalances for Alberta between 1975 and 1980; second, there is a comparison of graduate starting salaries between 1970 and 1980 in a sample of occupations, the salaries are also expressed as a ratio to average industrial earnings; third, enrolment shifts and applications are used as a measure of student responsive supply behavior; fourth, there is an adaptation of a sociological model, the world systems model, to explain labor market changes as one manifestation of general societal change in the late twentieth century. The findings reveal that during the time frame of the study there were large graduate shortages in many areas of specialization. The economic returns from a university education do not appear to be rapidly declining although students tend to be responding to the market through shifting enrolments and choices of field of study. The world systems model demonstrates much promise as a sociological tool for explaining social change.*

Increasing attention is being focused on the role of the university in Canadian society. The Senate of the University of Alberta, following in the footsteps of other Canadian institutions, recently published *The Report of the Commission on University Purpose* (1982). The social climate of the times seems to be providing the impetus for this current focus. The rapid, almost uncontrolled expansion of the higher education system is long past and the steadily increasing output of graduates from the universities during the 1970 decade has created new problems and concerns.

As long as universities in Canada remained elite institutions with a limited number of graduates entering a growing economy, employability was almost universal for graduates of all faculties. With the onset of the mass education philosophy in the late 1960s, there was a cumulative effect of larger graduating

---

Marilyn Westbury (PhD, Sociology of Education, University of Alberta) has taught school in Nova Scotia, British Columbia and Alberta; she is currently a lecturer in the Department of Educational Foundations at the University of Alberta. Her interests include labor market analysis and the education/labor market interface, teacher effectiveness, and women's studies.



classes. The latter, coupled with world recessionary trends in the 1970s, produced an anomaly—graduate unemployment. With the university no longer providing the ticket to guaranteed employment, the public began to question the purpose of the institution.

The debate about the role of the university is not new. For many years, proponents have touted the comparative benefits of a liberal arts education over more specialized, job-oriented programs in the university; that is, the cultural rewards of education versus the economic ones. The debate, however, has become less philosophical and more pragmatic. There is increasing public and governmental pressure on universities to become more responsive to the labor market. This pressure is intensified by the fact that it is the liberal arts graduates who are frequently in lowest demand in the labor market (Harvey, 1974, p. 149). It is easier to argue for the cultural benefits of a liberal arts degree when it is also a marketable commodity. However, the latter, rather than cultural benefits, may be the prime motivation for many students.

Although research in the area has been undertaken, the actual state of the labor market for university graduates has been more often an area of speculation than one of systematic study. Frequently repeated statements include: "Many university graduates are currently unemployed," "It is easier for technical institute graduates to find employment than it is university graduates," "There are declining financial returns from a university degree." Very few objective analyses are available to determine if these statements are based on fact. In a recessionary economy, it is important to look at the comparative position of graduates relative to other labor force entrants rather than at the isolated experiences of graduates today or fifteen years ago.

Public and governmental pressure seems to be suggesting that universities ought to be more responsive to labor market demand. Little is said about how institutions can achieve that goal. Presumably, the universities would have to make an effort to control enrolments to match the appropriate numbers of graduates in each area of specialization to projected labor market demand at the time of graduation three to ten years down the road. A more appropriate question might be, "*Can* the universities be more responsive to the labor market?" That is, do current methodology and social conditions permit accurate supply/demand predictions? Whether or not such a goal is desirable is a value judgment beyond the scope of this paper.

Recently completed research into the university education/employment relationship in Alberta is designed to explore systematically many of the above concerns. In addition, some comparisons with the graduate situation in the remainder of Canada are made. The areas investigated include job opportunities, wages, and student responsive supply behavior. A sociological theory of social change, the world systems model, is adopted as an explanation of labor market changes in the twentieth century.

### *Design of the Study*

The first thrust of the research is to measure, over time, the supply of university graduates from Alberta universities and to relate this supply to demand in the Alberta labor market for specific areas of specialization. In short, the purpose is to quantify labor market shortages or surpluses for university graduates in particular disciplines between 1975 and 1980. The intent of such supply/demand imbalances,

coupled with the growth in the managerial, professional and administrative segment of the labor force, is to determine the estimated job opportunities for graduates. Is it true that graduates have increased beyond the capacity for labor market absorption?

Actual and projected enrolments and graduations from Alberta universities constitute the supply (graduate imports are excluded). The Canadian Occupational Forecasting Program, COFOR (Canada Department of Employment and Immigration) is used as an estimate of demand. Regional variations in the model are prepared by Regional Economic Services of Employment and Immigration in conjunction with provincial government personnel. Two forecasts are used in this study, one prepared by Alberta Advanced Education and Manpower (COFOR, March, 1976) and a later one produced by Regional Economic Services (Canada Employment and Immigration, 1981). Both are estimates based on employment projections for 69 industries. Forecasts of industrial output, employment levels, and specific occupational requirements are linked to a medium term economic forecast for Canada with regional economic variations built into provincial forecasts. The estimates include net job creation and probable attrition through death and withdrawal for 496 occupational groups. Only those occupational groups which generally demand a university degree are segregated for this supply/demand study.

A second feature of the present research is the examination of starting salaries during the 1970 decade for a sample of occupations which require a degree. Does "wage-competition theory" (Thurow, 1979, p. 325) explain salaries paid to graduates? The theory posits that people with a particular package of skills compete in the marketplace with one another for wages which tend to rise or fall as skill shortages or skill surpluses become apparent. Wages are held to encourage or discourage individuals from seeking a particular skill package and as a result the labor market for each skill is kept close to balanced. Some researchers suggest that an increasing supply of university graduates, such as that which occurred in the last decade, would have the overall effect of reducing the salaries that employers are obliged to pay (Statistics Canada, 1978c). This thesis that there are declining economic returns from a university education is tested by comparing starting salaries of graduates from a sample of occupations between 1970 and 1980 with average industrial wages (industrial composite) during the same period. That is, a ratio of starting salaries for specific university graduates to average industrial wages over time is presented.

A third objective of this research is to determine if there is some evidence, in the form of altered enrolments, applications, or shifting faculty choices, of responsive supply behavior on the part of potential university students. This is an analysis of reciprocal or feedback aspects of the university education and employment interface which will ultimately have an impact on institutional policy.

A fourth and final thrust of the present study is the adaptation of a sociological tool, the world systems model, to provide an overall theoretical framework for explaining the social changes that have altered the university education/labor market relationship over the past few years. It is suggested that factors affecting the relationship between a university education and the labor market are best understood as smaller manifestations of global processes. That is, much of what is occurring in the Canadian labor market, for example, can be causally connected to the historical and economic circumstances of the world, its nations, and the regions within nations.



TABLE 1  
AVERAGE YEARLY ALBERTA SUPPLY AND DEMAND FOR MAJOR GROUPS  
OF MANAGERIAL AND PROFESSIONAL OCCUPATIONS, 1974-1979

|  | 1                          | 2                          | 3                                       | 4   | 5                        | 6  | 7                 | 8  | 9  |
|--|----------------------------|----------------------------|---|---|--------------------------|--|-------------------|--|--|
|  | Employment                 |                            | Total<br>Expansion<br>Demand<br>(2 - 1) | Total<br>Replacement<br>Demand<br>[Deaths]<br>[Withdrawals]<br>1975-1979 <sup>b</sup> | Total<br>1979<br>(3 + 4) | Demand<br>Percentage<br>of 1975<br>(5 ÷ 1) | Yearly<br>Average | Supply<br>Yearly<br>Average <sup>c</sup> | Yearly<br>Average<br>Supply-<br>Demand<br>Imbalance<br>(8 - 7) |
| CCDO<br>Code                                     | COFOR <sup>a</sup><br>1975 | COFOR <sup>b</sup><br>1979 |   |   |                          |  |                   |  |  |
| Occupational<br>Categories                       |                            |                            |   |   |                          |  |                   |  |  |
| 11 Managerial<br>Administrative                  | 32,030                     | 48,595                     | 16,565                                  | 5,949   | 22,514                   | 70.29                                      | 5,629             | 1,036                                    | -4,593   |
| 21 Natural Science<br>Engineering<br>Mathematics | 23,612                     | 40,265                     | 16,653                                  | 3,070   | 19,723                   | 83.52                                      | 4,931             | 1,004                                    | -3,927   |
| 23 Social Sciences                               | 7,131                      | 9,590                      | 2,459                                   | 962   | 3,421                    | 49.97                                      | 855               | 597                                      | - 258  |
| 25 Religion                                      | 1,847                      | 2,040                      | 193                                     | 363   | 556                      | 30.10                                      | 139               | 22                                       | - 117  |
| 27 Teaching                                      | 30,996                     | 44,330                     | 13,334                                  | 4,273   | 17,607                   | 56.80                                      | 4,402             | 1,859                                    | -2,543   |
| 31 Medicine and<br>Health                        | 36,614                     | 44,005                     | 7,391                                   | 4,636   | 12,027                   | 32.85                                      | 3,007             | 491                                      | -2,516   |
| 33 Arts and<br>Recreation                        | 6,033                      | 9,405                      | 3,372                                   | 654   | 4,026                    | 66.73                                      | 1,007             | 151                                      | - 856  |
| TOTAL ALL CATEGORIES                             | 138,263                    | 198,230                    | 59,967                                  | 19,917  | 79,874                   |  | 19,970            | 5,160                                    | -14,810  |

Sources of Data:

<sup>a</sup>Alberta Advanced Education and Manpower, Planning Secretariat, Summary of Canadian Occupational Forecast (COFOR) Model Results for Alberta 1975-1982. (Edmonton: March, 1976)

<sup>b</sup>Canada Employment and Immigration, Occupational Requirements to 1985: Canadian Occupational Forecasting Program (COFOR). (Ottawa: January, 1981)

<sup>c</sup>Statistics Canada, Universities: Enrolments and Degrees. Catalogue No. 81-204, 1975-1979. Graduates available to the labor force plus reserve re-entrants.

### *Findings and Conclusions*

The supply-demand analysis for Alberta produced some extensive supply shortages for the major groups of managerial-professional occupations segregated in the *Canadian Classification and Dictionary of Occupations* (1971). After attrition in supply (a fairly predictable number do not, for various reasons, seek employment in Alberta within one month of graduation), there were shortages in the following seven major groups of managerial professional occupations: managerial, administrative; natural sciences, engineering and mathematics; social sciences; religion; teaching; medicine and health; arts and recreation. In all of the former groups, provincial graduate supply was inadequate to meet Alberta demand between 1975 and 1979. In a sample of fourteen individual occupations, there were four in which there was a surplus of graduates (See Tables 1 and 2).

The data point out the extent to which graduate imports to Alberta filled manpower vacancies between 1975 and 1980. Employment in managerial, professional and administrative occupations in Alberta grew from 161,000 in 1975 to 219,000 at year end 1979 (Statistics Canada, 1975-1979d). This increase far exceeds the total degrees granted (37,501) in Alberta over the same period of time (Statistics Canada, 1975-1979d). While some of the managerial, professional, and administrative occupations are filled by non-degree holders, the vast majority of recently filled positions carry degree prerequisites. Moreover, related literature indicates that roughly 30 percent of graduate output does not move into managerial/professional type jobs (Statistics Canada, 1981). With a retention rate of 63 percent, approximately 23,600 Alberta degree graduates were potential competitors for 58,000 Alberta managerial, professional, and administrative positions between 1975 and 1979. With the possible movement of roughly 30 percent of the graduates into sales and other nonprofessional/managerial/administrative occupations, the potential number of Alberta competitors is further reduced to approximately 16,500 (See Table 3).

The growth in the appropriate segment of the labor force suggests a number of job opportunities for highly qualified applicants in Alberta in the time period segregated. It also illustrates the mobility of university graduates in Canada. Unemployment among university graduates in the province remained consistently low at between two and three percent, often one-third as high as general unemployment. In addition, unemployment rates for university graduates were consistently below those of college graduates and others with less training (Statistics Canada, 1975-1980b). University graduates were more marketable than other labor force entrants in Alberta and Canada. The large migration to the province suggests that employment opportunities were competitive with those elsewhere in Canada. Alberta had a higher rate of absorption for more highly qualified manpower than for any other segment of the labor force (Statistics Canada, 1975-1979b). Although a university degree did not guarantee employment, labor force entrants were still better off with a degree than without.

The results of the data on salaries of university graduates (although the sample of five occupations was too small to generalize conclusions) showed little support for the thesis that an increasing supply of university graduates was reducing the salaries that employers are obliged to pay. Starting salaries for the five occupations, relative to average industrial wages in Alberta, did not show a consistent drop over the decade. Both the industrial composite and the graduate starting salaries increased with some specialization variation (See Table 4). Also, there was little



support for wage competition theory. That is, wages do not always appear to rise and fall as skill shortages or skill surpluses became apparent. Nor do low wages in a field consistently discourage potential students from entering that specialization. Finally, fields with the greatest shortages do not necessarily pay the highest starting wages in the labor market. In summary, it would appear that many factors including unions, professional associations, traditions of wage differentials, training costs, and technical progress all influence the wage associated with each job (Thurow, 1979, p. 330).

TABLE 2

AVERAGE YEARLY ALBERTA OCCUPATIONAL SUPPLY AND DEMAND  
FOR SELECTED OCCUPATIONS, 1975-1979

| CCDO<br>Code  | Occupations   | 1  | 2  | Average Yearly<br>Supply-Demand<br>Imbalance<br>1975-1979<br>(2 - 1) |
|---------------|---|--|--|--|
|               |   | Requirements <sup>a</sup><br>Average Yearly<br>Demand<br>1975-1979 | Supply <sup>b</sup><br>Average Yearly<br>Degree Output<br>With Appropriate<br>Specialization |  |
| 1171          | Accountants<br>Auditors<br>Financial Officers                   | 915  | 407  | -508   |
| 2111          | Chemists  | 38   | 34   | - 4  |
| 2112          | Geologists  | 196  | 65   | -131   |
| 2113          | Physicists  | 3  | 27   | + 24   |
| 2142-<br>2154 | All Engineers   | 449  | 292  | -157   |
| 2181          | Mathematicians<br>Statisticians<br>Actuaries                    | 16   | 41   | + 25   |
| 2183          | Systems Analysts<br>Computer Programmers<br>Related Occupations | 112  | 52   | - 60   |
| 2311          | Economists  | 24   | 70   | + 39   |
| 2313          | Sociologists  | 1  | 35   | + 34   |
| 2341-<br>2343 | Judges<br>Lawyers   | 184  | 103  | - 81 <sup>c</sup>  |
| 3111          | Physicians and<br>Surgeons                                      | 410  | 175  | -235 <sup>c</sup>  |
| 3113          | Dentists  | 111  | 29   | - 82 <sup>c</sup>  |
| 3151          | Pharmacists   | 88   | 57   | - 31   |
| 3352          | Writers and<br>Editors<br>Publication                           | 103  | 44   | - 59   |

<sup>a</sup>COFOR 1975-1982, averaged yearly.

<sup>b</sup>Supply is degree output subject to .63 retention ratio. Between 1970 and 1980, approximately 37% of Alberta university graduates were not seeking employment in Alberta one month after graduation.

<sup>c</sup>Likely to be a slightly overstated shortage because of a probable lower attrition rate from graduation to the labor force.

TABLE 3

ANNUAL GROWTH OF EMPLOYMENT IN MANAGERIAL, PROFESSIONAL  
AND TECHNICAL OCCUPATIONS AND LABOR FORCE ENTRANTS WITH  
POSTSECONDARY EDUCATION, CANADA AND ALBERTA

| Year                           | 1   |                  | 2                  |       | 3   |       | 4   |       | 5   |       | 6  |       |
|--------------------------------|---|------------------|--------------------|-------|---|-------|---|-------|---|-------|--|-------|
|                                | Employment in<br>Managerial,<br>Professional<br>Administrative<br>Occupations |                  | Annual<br>Increase |       | Annual %<br>Increase in<br>Employment<br>( $\frac{2}{1} \times 100$ ) |       | Number of<br>Labor Force<br>Entrants with<br>University<br>Degree |       | Annual %<br>Increase of<br>Labor Force<br>Entrants with<br>University<br>Degree |       | Annual %<br>Increase of<br>Labor Force<br>Entrants with<br>Postsecondary<br>Graduation |       |
|                                | Can.<br>(000's)   | Alta.<br>(000's) | Can.               | Alta. | Can.  | Alta. | Can.  | Alta. | Can.  | Alta. | Can.   | Alta. |
| 1975                           | 2008  | 161              | -                  | -     | -   | -     | n/a   | n/a   | -   | -     | -  | -     |
| 1976                           | 2095  | 176              | 87                 | 15    | 4.3   | 9.3   | 45  | 17    | 4.9   | 21.8  | -4.3   | 1.0   |
| 1977                           | 2161  | 183              | 66                 | 7     | 3.2   | 4.0   | 58  | 2     | 6.0   | 2.1   | 2.4  | 3.5   |
| 1978                           | 2255  | 215              | 94                 | 32    | 4.4   | 17.5  | 54  | 13    | 5.3   | 13.4  | 4.4  | 5.3   |
| 1979                           | 2373  | 219              | 116                | 4     | 5.1   | 1.9   | 73  | 6     | 6.8   | 5.5   | 1.4  | 2.7   |
| 1980                           | 2440  | 231              | 67                 | 12    | 2.8   | 5.5   | 75  | 2     | 6.5   | 1.7   | 6.1  | 6.6   |
| Annual Average<br>Over 5 Years | -   | -                | 86                 | 14    | 4.0   | 7.6   | 61  | 8     | 5.9   | 8.9   | 2.0  | 3.8   |

Sources:  
Statistics Canada, The Labor Force, Catalogue #71-001. Annual averages 1975-1980.  
Statistics Canada, Out of School -- Into the Labor Force, Catalogue #81-570E, p. 155.

TABLE 4

RATIO OF ANNUAL STARTING SALARIES<sup>a</sup> IN INDUSTRY  
FOR FIVE OCCUPATIONS TO AVERAGE EARNINGS IN INDUSTRY<sup>b</sup>  
CANADA, EDMONTON, 1970-1980

|      | Engineer |          | Computer Systems<br>Administrator |          | Financial<br>Administrator |          | Physical<br>Scientist |          | Economist<br>Statistician |          |
|------|----------|----------|-----------------------------------|----------|----------------------------|----------|-----------------------|----------|---------------------------|----------|
|      | Canada   | Edmonton | Canada                            | Edmonton | Canada                     | Edmonton | Canada                | Edmonton | Canada                    | Edmonton |
| 1970 | 1.32     | 1.33     | 1.30                              | 1.43     | 1.35                       | 1.39     | n/a                   | n/a      | 1.31                      | n/a      |
| 1971 | 1.26     | 1.33     | 1.30                              | 1.45     | 1.36                       | 1.48     | n/a                   | n/a      | 1.60                      | n/a      |
| 1972 | 1.22     | 1.27     | 1.28                              | 1.62     | 1.48                       | 1.43     | n/a                   | n/a      | 1.56                      | n/a      |
| 1973 | 1.18     | 1.21     | 1.24                              | 1.16     | 1.36                       | 1.39     | n/a                   | n/a      | 1.48                      | n/a      |
| 1974 | 1.17     | 1.22     | 1.24                              | 1.18     | 1.41                       | 1.43     | 1.36                  | n/a      | 1.53                      | 1.66     |
| 1975 | 1.23     | 1.27     | 1.21                              | 1.05     | 1.39                       | 1.35     | 1.38                  | n/a      | 1.55                      | 1.79     |
| 1976 | 1.20     | 1.22     | 1.20                              | 1.17     | 1.35                       | 1.47     | 1.34                  | n/a      | 1.54                      | 1.85     |
| 1977 | 1.18     | 1.17     | 1.19                              | 1.10     | 1.32                       | 1.41     | 1.29                  | n/a      | 1.48                      | 1.62     |
| 1978 | 1.18     | 1.25     | 1.19                              | 1.24     | 1.38                       | 1.49     | 1.38                  | 1.37     | 1.49                      | n/a      |
| 1979 | 1.18     | 1.20     | 1.20                              | 1.18     | 1.38                       | 1.31     | 1.38                  | n/a      | 1.41                      | n/a      |
| 1980 | 1.22     | 1.22     | 1.22                              | 1.14     | 1.38                       | 1.34     | 1.42                  | n/a      | 1.41                      | n/a      |

Sources:  
<sup>a</sup>Pay Research Bureau, Canada Public Service Staff Relations Board. Data compiled from computer tapes made available by N. B. Allen, Project Coordinator, Ottawa, 1981. Salary rate measures are surveyed every August, except for 1970 when they were surveyed in October.  
<sup>b</sup>Average weekly earnings (Industrial Composite) X 52. Statistics Canada, Employment and Hours, Catalogue No. 72-002.

\* Mean starting salaries + average earnings in industry.



Related literature does suggest that starting salaries in some fields of specialization may be deteriorating relative to general wages. Anticipated recruiting rates of pay for general arts (not honors) graduates in *Canada* did deteriorate between 1975 and 1980 relative to average earnings in industry (Pay Research Bureau, 1975-1980). In many other faculties, however, relative wages improved or remained the same. In summary, there is no consistent trend.

Two additional factors are important when discussing graduate starting salaries. First, starting wages may carry less weight with students than potential wages ten years hence. Second, research indicates that *average* salaries for graduates tend to increase as levels of education increase (Statistics Canada, 1981, p. 251). That is, Ph.D. graduates generally earn higher salaries than Master's graduates. It is also true that university graduate wages are, on the average, higher than college graduates. While some university graduates may be earning salaries below their expectations, most are still comparatively better off than non-university labor force entrants. The findings exclude educational costs and foregone earnings.

The third objective of the study was to examine evidence that would indicate responsive supply behavior on the part of students. Freeman concluded that "there is a substantial and growing body of evidence that young persons . . . are highly sensitive in their educational and career decisions to the state of the labor market" (1976, p. 53). It is assumed that two barometers of responsive supply behavior are student enrolments and the distribution of students across various fields of specialization. In general, the study provided only moderate support for Freeman's conclusion. Some general trends indicate that student choices (as measured by number of applications) are shifting to more traditionally marketable fields such as medicine and law, and away from fields of declining marketability such as Ph.D. programs. There are also increases in the number of qualified applicants to business and commerce schools in Alberta, an area with increased labor market demand. On the other hand, some fields, such as sociology, physics, and chemistry, had fairly consistent enrolments in spite of an estimated surplus of graduates in the labor market. Although teachers were in high demand in Alberta, enrolments dropped substantially. The education potentials may have been responding to media emanating from outside of Alberta where an oversupply of teachers was apparent. There are several additional factors which may possibly mitigate responsive supply behavior. Graduate imports probably altered the stated supply/demand imbalances in most fields and perhaps students are more responsive to the real labor market than this research is able to reveal. In addition, up-to-date information on the labor market is not always available to students nor are all students willing to abandon a chosen field should job opportunities appear scarce. Finally, the 1970 decade was a period of declining political, social, and economic support for universities. Funding cutbacks, tighter admissions policies, additional quota faculties and a more negative climate of opinion than had been prevalent in the free-spending, white collar promotional days of the 1960s quite possibly dampened student response.

The fourth thrust of the present research is the adaptation of a theoretical model for explaining social change processes and the labor market shifts that are manifestations of larger social and economic conditions. The world systems model, introduced by Wallerstein (1974) and effectively applied to the twentieth century by Chirot (1977), is a most appropriate vehicle for predicting and explaining modern events in what has become a global economic network. One of the key events which emphasized this global interdependency was the OPEC oil cartel in 1973. The impact on western industrialized nations was dramatic. Canada, as a net

importer of oil, was faced with a spiralling energy bill because of the inability to immediately reduce either oil purchases or consumption. Since energy increases push up the consumer price index while outflows of money depress the economy, the net result is increasing rates of unemployment coexisting with high inflation. On the other hand, the OPEC cartel produced an economic boom in Alberta, a major producer of oil, and employment opportunities within the province grew. A particularly large demand for highly qualified technological and economic specialists with training relevant to the industrial focus of Alberta was apparent.

Overall, the world systems model is a useful tool for explaining social change. It outlines the presence of regional-national hostility ("regional nationalisms") because of inequalities in power and wealth, it accentuates the fact that world economic events quickly reverberate to the local level and are mirrored by labor market demand, and it points to the presence of both economic and cultural dominance in the world. The abruptness with which labor markets do change is very apparent in the turnaround in Alberta in 1981. The much lower demand and higher unemployment in 1982 are primarily results of two world events: reduced oil and gas consumption in both Canada and the world and a glut of oil on world markets holding prices at a lower than anticipated level.

The economic recession in 1981-82 has finally brought national trends closer to the Alberta scene. Alberta university participation rates had dropped below most of the provinces in Canada during the late 1970s when job opportunities appeared to be mushrooming in the province. Increased enrolments during the 1980s, along with an aging student population, appear to confirm that job availability lures students into the labor force earlier while low demand forces many unemployed back to the university to improve their competitive position.

### *Implications*

A question posed earlier was, "Should the universities be more responsive to labor market demand?" Of course, the answer to this question rests on another, "Can the universities be more responsive to labor market demand?" The research described here suggests that attempts by institutions to match graduate output to labor market demand for particular fields of study would present problems. However, the issue is more complex than these simple questions suggest and is ultimately reduced to whether or not it is more desirable to use a limited manpower projection system or none at all in institutional planning. After presenting some of the practical problems associated with occupational supply and demand projections, a list of how and why the latter may nevertheless be useful to universities will conclude this paper.

Those using projection data should be well informed about the uncertainties and limitations of the data. In the first place, one of the biggest difficulties is the open nature of the system. Universities are provincially governed yet graduates are a very mobile group. For example, while Alberta universities graduated a shortage of highly qualified workers in most fields of study between 1975 and 1979, job vacancies were frequently filled by migration, primarily interprovincial, to the province. It would be very difficult to predict interprovincial migration several years in advance. Thus, although it is currently feasible to obtain a rough estimate of labor market requirements for short and medium term forecasting, it is quite another matter to estimate the numbers from outside the province who will choose to compete for those vacancies. As a case in point, between 1975-76 and 1979-80, a



fairly consistent annual number of between 915 and 1,171 interim certificates were granted to teacher imports to Alberta (Loken, 1980, p. 71). After Alberta Education introduced a portability clause in January, 1982, whereby out-of-province credentials were accepted for certification, there were 3,124 interim certificates granted to teacher imports in the first ten months of 1982 (Alberta Education, November 22, 1982).

A second formidable problem in attempting to match supply to demand for university graduates is the notion of lag. For the highest degrees, long term forecasts ten years hence are required but general experience has proven long-term forecasting to be highly inaccurate. The same problem of time lag invades student responsive supply behavior. The study described here suggests that the diminished marketability for certain levels of specialization, such as the Ph.D. degree, appears to have had the effect of increasing applications to historically marketable fields such as medicine and law. While limitations on enrolments are maintained through quotas, demand in the latter fields is diminishing because of a consistent, moderate provincial output over a number of years combined with a surplus of graduates in some provinces of Canada. A cyclical pattern of shortages and surpluses emerges in some fields of study. Freeman (1971) illustrates that this is particularly noticeable in the field of engineering where the existence of a shortage of personnel and plentiful job opportunities in a particular area of specialization attracts a greater freshman enrolment in the subsequent year which, in turn, causes a surplus after the normal lag period to graduation. The process repeats itself when the graduate surplus discourages freshman enrolment and another shortage ensues after the appropriate four year lag. For the Ph.D. degree, lag time is so extensive that long term shortages have, in the past, been filled by foreign recruitment. In summary, it is difficult for institutions to produce accurate long-range forecasts (which rely on economic forecasts) for determining Ph.D. demand, for example, ten years in advance. In addition, student response always lags behind demand for the length of the particular training required.

Another problem area in balancing university supply to labor market demand is the possibility of an abrupt change in the economy. Market demand for all graduates may shift radically during the training period. The selection of the world systems model of social change is an attempt to illustrate national vulnerability and global interdependency in the late twentieth century. On the national as well as the Alberta front, economic waves of prosperity and recession are neither predictable nor controlled by provinces or nations. Moreover, the model suggests that the capitalist world is now entering a contraction period and that rapid economic expansion is the thing of the past. This latter problem alone will have a profound impact on the job prospects of university graduates, already a steadily increasing segment of the population.

It is apparent that there are practical problems in having universities respond to manpower projections. Nevertheless, good projections could be useful to universities as one of many administrative tools. Manpower planning is frequently criticized on the philosophical basis that it is neither the role nor the responsibility of the university to deflect people from preferred career paths. Current practices, such as the use of enrolment quotas in an increasing number of faculties, mean that universities are already engaged in a deflecting and sorting process which may or may not be largely labor market oriented. The rationale for the latter kind of decision making is rarely made public.



For several reasons universities could benefit from manpower projections, albeit imperfect ones. First, outside pressure to use occupational supply/demand imbalances is increasing and the data are acceptable to a public who have otherwise been uninformed about the logic behind higher educational alterations. Second, students frequently request information about employment opportunities on which to base career decisions. Third, universities can use the information for program modification and increased flexibility in an era of rapid change.

By incorporating manpower projections into other planning approaches, universities need not abandon that very important traditional role: to graduate students with the ability to solve problems, analyze, communicate effectively, and who are independent learners capable of adapting to new situations, jobs, and challenges as they come along.

### References

- Alberta Advanced Education and Manpower. Planning Secretariat. (1976, March). *Summary of Canadian Occupational Forecast (COFOR) model results for Alberta 1975-1982*. Edmonton.
- Alberta Education. (1982, November 22). Copy of a letter from L. E. Rappel, Registrar. Effect of portability upon certification numbers. Edmonton.
- Canada Employment and Immigration. (1971). *Canadian classification and dictionary of occupations*. Ottawa: Information Canada. *Annual guide, 1978-79*.
- Canada Employment and Immigration. (1981, January). *Occupational requirements to 1985: Canadian Occupational Forecasting Program (COFOR)*. Ottawa.
- Chiot, D. (1977). *Social change in the twentieth century*. New York: Harcourt Brace Jovanovich, Inc.
- Freeman, R. (1976). *The market for college-trained manpower*. Cambridge, MA: Harvard University Press.
- Harvey, E. (1974). *Educational systems and the labour market*. Don Mills, Ontario: Longman Canada Limited.
- Loken, G. (Principal Investigator) (1980). *Projecting K-12 teacher supply and demand in Alberta*. Edmonton: Alberta Education.
- Pay Research Bureau. Canada Public Service Staff Relations Board. (1975-1980). *Anticipated recruiting rates for university and community college graduates*. Ottawa: Annual.
- Statistics Canada. *Employment earnings and hours*. Cat. No. 72-002. Ottawa: Annual. (a)
- Statistics Canada. (1975-1980). *The labour force*. Cat. No. 71-001. Ottawa: Annual Averages. (b)
- Statistics Canada. (1978). *Out of school—into the labour force*. Cat. No. 81-570E. Ottawa. (c)
- Statistics Canada. (1975-1979). *Universities: Enrolments and degrees*. Cat. No. 81-204. Ottawa. (d)
- Statistics Canada. Education, Science and Culture Division. (1981, March). *Job market reality for postsecondary graduates*. Cat. No. 81-572E. Ottawa. (e)
- The Senate, University of Alberta. (1982, April). *Report of the Commission on University Purpose*.
- Thurrow, L. C. (1979). Education and economic equality. In J. Karabel & A. H. Halsey (Eds.), *Power and ideology in education*. New York: Oxford University Press.
- Wallerstein, I. (1974). *The modern world-system: Capitalist agriculture and the origins of the European world economy in the sixteenth century*. New York: Academic Press.

GRACE M. H. GAYLE

University of Ottawa

## Second Language Teaching Styles: Making or Breaking Attitudes

*The possibility that direct and/or indirect relationships between attitudes and second language learning may be differentially affected by specific second language teaching styles was examined. Second language teaching styles refer to the sustained configuration of percentage frequencies on categories of discourse functions which are considered to be accurately descriptive of classroom oral events. The Language Teaching Record Scheme (LTRS) was used to code classroom interaction for a total of 3 random visits for each of 14 teachers taken from the Intensive Programme. The cluster analysis technique yielded four distinct styles of which three were found to be relevant to attitudes. Analyses of covariance indicated that the hypothesized interaction of teaching style with attitudes towards language learning was supported. Significant interaction effects ( $p < .10$  or better) between specific teaching styles and specific categories of students measured on Attitudes towards French Canadians, Attitudes towards Learning French, and Desire to Learn French were found on adjusted group means for aural comprehension scores. Pretreatment scores on aural comprehension were used as the covariate. Significant main effects ( $p < .05$ ) were also found for distinct teaching styles in the same programme.*

The direct association between attitudes and second language achievement is now firmly established (Burstall, 1975; Chastain, 1976; Gardner & Lambert, 1972; Mueller & Miller, 1970; Tucker, Hamayan & Genesee, 1975). On a more basic level, attitudes are also seen as a complex support system for motivation in second language learning (Gardner, Glikson & Smythe, 1978). From these findings, it may be concluded that the relationship between attitudes and achievement in second language learning is both direct and indirect. Recently, a question of a very different but not unrelated order has begun to receive attention. Research on second language teaching styles (Gayle, 1982b) has attempted to explore the

---

Grace M. H. Gayle (PhD, Ottawa) is Assistant Professor, Faculty of Education, University of Ottawa. She has also served as a Research Officer at the Ontario Institute for Studies in Education. Her major research and teaching interests have been in the area of learning with emphasis on psycholinguistics. Since 1978, she has published on Second Language Teaching Styles; publications include a book which details a methodology in identifying styles.

importance of this classroom variable to learning. This research was inspired in part by conjectures and hypotheses relevant to the broader educational setting on the possibility of student/treatment interactions (Cronbach, 1957; Lewin, 1935) and, more specifically, by interactions discussed in the general literature between student characteristics and teaching style (Canfield, 1978; Flanders, 1970; Hunt, 1971; Myers, 1980; Pervin, 1968). Gayle found that second language teaching styles do make a difference to the relationship between specific student aptitudes and student learning. This led to the interesting speculation that, even if considered incapable of eradicating negative attitudes entirely, specific second language teaching styles may be sufficiently potent to break the behavioural (withdrawal) component (Katz & Stotland, 1959) which accompanies the affective ingredient of a negative attitude.

Second language teaching styles have been defined (Gayle, 1982a, 1982b) as the sustained configuration of techniques or teacher characteristics identified on specific dimensions of second language teaching—that is, configurations which pertain to discrete behavioural or characterological categories combined to form well defined teaching *patterns*. Previous to this, teaching techniques had been discussed extensively with reference to possible outcomes in both general content classrooms and the second language classroom (Fanselow, 1977; Mackey, 1965; Orme, 1978). However, the one to one study of specific techniques and learning outcomes or of global, unanalyzed strategies and learning outcomes can hardly lead to a complete understanding of what brings learning about. Gayle (1980a, 1982b) suggests that a dynamic and complex process such as second language learning can only be understood if the conditions of learning provided by the teacher are perceived as a systematic gestalt of contributing but clearly identifiable parts. With this conviction it appears that there is further need for the study of the effects of second language teaching style when the dimension(s) of concern have been analyzed in terms of a set of exhaustive but discrete categories.

The concept of teaching style should also offer a better explanation of the nature of second language classroom interaction than the traditional reference to method since styles are specified in retrospect after intensive observation of behaviours which actually occurred. That is, teaching styles are not prescriptive but descriptive. Specifically, teaching style would avoid the type of confusion attributed by Mackey to past use of the term, method, where it was not the presentation done by the teacher that was being analyzed, but rather that done by the method through its texts, films, and recordings (Mackey, 1967). Teaching style brings the researcher into direct contact with the teacher and with classroom events.

In describing teaching styles or the composite of classroom events enacted or allowed by the teacher, care must be taken to collect data by the use of observation schemes which are based on discrete non-redundant categories. That is, a valid conclusion cannot be reached regarding the relationship of teaching style to learning outcome if inter-coder reliability and/or the validity of the instrument has been prevented by the presence of ambiguous or overlapping categories.

Previous exploratory research (Gayle, 1982b) implementing the Language Teaching Record Scheme (LTRS) (Gayle, 1980a) identified a number of second language teaching styles based on percentage frequencies for discourse functions. The LTRS is a second language classroom observation instrument which attempts to analyze teaching on a number of dimensions using categories within each dimension which are exhaustive but discrete. On this basis it facilitates the study of



student outcomes which may be interpreted as the result of combinations of teacher behaviours. Discourse functions refer to only one dimension of teaching analyzed by the LTRS and describe the *pedagogical* function of the oral behaviour rather than its linguistic function. A statistical cluster analysis of teachers was derived from data for individual teachers using percentage frequencies for all discourse functions simultaneously. These data were pertinent to both teacher talk and student talk. Four distinct clusters associated with four distinct styles emerged. Three of these styles, identified here as H, J, and L, are displayed in Table 1. Figures shown in Table 1 represent total percentage frequencies for all teachers in their respective clusters. Asterisks in Table 1 highlight those categories which receive comparatively greater emphasis than their counterparts in the contrasting style. Table 2 gives a relatively detailed definition of each category.<sup>1</sup>

The concept of teaching style adheres to the notion that a particular teaching style is a relatively stable phenomenon within a cluster of teachers and tends to exhibit some features which may be referred to as dominant since they quantitatively overshadow the rest. The findings recorded in Table 1 illustrate that this dominance is often complemented by another type of preponderance which is the direct outcome of comparisons between styles as the quantitative strength of the categories in one style is compared with that of their counterparts in another style on the horizontal plane.

Categories in Table 1 have been arranged in order to juxtapose those discourse functions which appear, on the basis of previous research (Gayle, 1982b), to be associated logically and/or empirically. These groups have been regarded as strategies which can be used to facilitate discussion and comparison of styles as well as to suggest explanation of their effects. Six strategies have been identified: *Basic or Stock Strategy*, *Language Exposure*, *Independent Analysis and Use of Language (Enterprise)*, *Mechanical Use of Language*, *Direction*, and the *Verbal Processing* strategy.

The *Basic or Stock* strategy is comprised of the single discourse function, specific information, which is regarded as the “filler” category for any style since it serves as the backbone for all classroom interaction in the Intensive Programme.<sup>2</sup> Its frequency tends to be obviously reduced when other discourse functions are more heavily used.

The term *Language Exposure* is intended to convey the idea that, with use of these categories, students are allowed more extensive experience with the language beyond the minimal question/answer paradigm. This strategy covers the categories general information, reading, writing, illustration and rephrasal.

The strategy *Independent Analysis and Use of Language* describes the teacher’s or student’s intense readiness to use the language to the point of being willing to initiate a *change* (digression) in the subject matter in order to prolong the discussion. Occurrence or encouragement of this extreme type of language readiness is measured by the single category, initiation. The strategy *Independent Analysis and Use of Language* also includes the use of the discourse function, implicit correction, which calls on the student to analyze language independently and extrapolate the essence of the rule. In summary, the student must be confident enough to use language creatively and extensively and to make linguistic inferences as a budding entrepreneur would.

TABLE 1  
TEACHING STYLES IDENTIFIED FROM A CLUSTER ANALYSIS<sup>a</sup>  
OF TEACHERS AND BASED ON TOTAL PERCENTAGE FREQUENCIES  
FOR DISCRETE DISCOURSE FUNCTIONS

| Strategy  | Discourse Functions          | H<br>Lateral | J<br>Balanced | H<br>Lateral | L<br>Linear | J<br>Balanced | L<br>Linear |
|---|------------------------------|--------------|---------------|--------------|-------------|---------------|-------------|
| BASIC   | Specific Information         | 151.30*      | 135.73        | 151.30       | 157.75*     | 135.73        | 157.75*     |
| LANGUAGE EXPOSURE                                   | General Information          | 6.73         | 8.37*         | 6.73         | 10.41*      | 8.37          | 10.41*      |
|   | Reading                      | 57.21*       | 40.32         | 57.21*       | 14.60       | 40.32*        | 14.60       |
|   | Writing                      | 0.00         | 13.64*        | 0.00         | 0.00        | 13.64*        | 0.00        |
|   | Illustration                 | 0.33         | 0.57*         | 0.33         | 1.37*       | 0.57          | 1.37*       |
|   | Rephrasal                    | 7.21*        | 2.20          | 7.21         | 7.22*       | 2.20          | 7.22*       |
|   |                              | 71.48        | 65.10         | 71.48        | 33.60       | 65.10         | 33.60       |
| INDEPENDENT ANALYSIS & USE OF LANGUAGE (ENTERPRISE) | Initiation                   | 0.41*        | 0.37          | 0.41*        | 0.00        | 0.37*         | 0.00        |
|   | Implicit Correction          | 6.23*        | 2.23          | 6.23*        | 5.03        | 2.23          | 5.03*       |
|   |                              | 6.64         | 2.60          | 6.64         | 5.03        | 2.60          | 5.03        |
| MECHANICAL USE OF LANGUAGE                          | Complete Repetition          | 8.73         | 23.31*        | 8.73         | 43.54*      | 23.31         | 43.54*      |
|   | Partial Repetition           | 2.79*        | 2.65          | 2.79         | 5.85*       | 2.65          | 5.85*       |
|   | Expansion to Full Sentence   | 2.05         | 4.15*         | 2.05         | 9.16*       | 4.15          | 9.16*       |
|   |                              | 13.57        | 30.11         | 13.57        | 58.55       | 30.11         | 58.55       |
| DIRECTION   | Acceptance                   | 25.12*       | 20.38         | 25.12*       | 15.64       | 20.38*        | 15.64       |
|   | Rejection                    | 2.29         | 4.38*         | 2.29*        | 1.73        | 4.38*         | 1.73        |
|   | Directive                    | 11.98*       | 8.75          | 11.98*       | 3.80        | 8.75*         | 3.80        |
|   |                              | 39.39        | 33.51         | 39.39        | 21.17       | 33.51         | 21.17       |
| VERBAL PROCESSING                                   | Explicit Correction          | 9.91*        | 8.41          | 9.91*        | 7.40        | 8.41*         | 7.40        |
|   | Indication of Possible Error | 1.52         | 3.46*         | 1.52         | 2.20*       | 3.46*         | 2.20        |
|   | Clarification                | 1.96         | 7.62*         | 1.96*        | 0.14        | 7.62*         | 0.14        |
|   | Grammar Clue                 | 0.00         | 1.89*         | 0.00         | 0.59*       | 1.89*         | 0.59        |
|   | Spec. Info. (clue)           | 2.73         | 7.51*         | 2.73         | 11.65*      | 7.51          | 11.65*      |
|   | Gen. Info. (clue)            | 0.16         | 0.69*         | 0.16         | 0.17*       | 0.69*         | 0.17        |
|   | Illustration (clue)          | 0.00         | 0.49*         | 0.00         | 0.00        | 0.49*         | 0.00        |
|   | Rephrasal (clue)             | 0.27         | 2.29*         | 0.27         | 1.03*       | 2.29*         | 1.03        |
|   | Cross-lingual Clue           | 0.41*        | 0.31          | 0.41*        | 0.00        | 0.31*         | 0.00        |
|   | Translation                  | 0.41         | 0.57*         | 0.41         | 0.50*       | 0.57*         | 0.50        |
|   |                              | 17.37        | 33.24         | 17.37        | 23.68       | 33.24         | 23.68       |

*Note.* Asterisks highlight those categories in a style which have higher frequencies than their counterparts in the comparison style.

<sup>a</sup>Each group chosen from the cluster analysis for study was comprised of three teachers.

TABLE 2  
DESCRIPTION OF DISCOURSE FUNCTIONS

| Discourse Functions          | Description  |
|------------------------------|--|
| Specific Information         | The use of close-ended questions which elicit the single correct answer. Responses which conform to this type of question and therefore tend to be stereotypical.  |
| General Information          | The use of open-ended questions. Creative use of language including the presentation of comments, opinions, arguments, personal experiences or fictional statements which have not been rehearsed in the classroom.              |
| Reading                      | The act of reading either aloud or silently a word, sentence or series of sentences in order to practise the skill or to focus attention on the subject matter of subsequent interaction.  |
| Writing                      | The request for or execution of writing during oral interaction.   |
| Illustration                 | Any example of a previous rule or generalization (regarding content or structure).   |
| Rephrasal                    | Expression of a statement or question in different words without the basic meaning being changed and without adding information. This category includes a summary of a previous passage or statement.                            |
| Initiation                   | Extreme readiness to utilize language freely and extensively as seen when the speaker continues to speak, <u>after</u> his response has been given, by voluntarily <u>changing the subject matter</u> .                          |
| Implicit Correction          | Substitution of the error with the correct form without drawing attention to the nature or location of the error.  |
| Complete Repetition          | Use or elicitation of <u>verbatim</u> repetition.  |
| Partial Repetition           | Use of elicitation of repetition of only the <u>essential</u> part of a preceding response for reinforcement or confirmation.  |
| Expansion to Full Sentence   | Teacher insistence on the use of the grammatically complete sentence. Emphasis on form rather than essential content.  |
| Acceptance                   | Positive evaluation (agreement with) the content or form of a previous response.   |
| Rejection                    | Negative evaluation (disagreement with) the content or form of a previous response.  |
| Directive                    | A command or request for a response which is nonverbal (an action or activity).  |
| Explicit Correction          | Indicating exactly <u>what</u> was wrong.  |
| Indication of Possible Error | Correction which is purely suggestive and offers no specific help to the student; for example, "Est-ce qu'on dit ça?"  |
| Clarification                | The <u>reason</u> for the correctness/incorrectness of the linguistic form or content of the previous response. The clarification is not a clue (i.e., it is given <u>after</u> the response).                                   |
| Grammar Clue                 | Reference to or citation of the grammatical features of the language being learned in order to help the student arrive at the correct answer.  |
| Specific Information (clue)  | Communication of specific information required by the student to arrive at <u>the</u> correct answer.  |
| General Information (clue)   | Communication of general information which may be used by the student to arrive at any acceptable answer.  |
| Illustration (clue)          | Illustrations which enable the student, by induction or analogy, to find an answer.  |
| Rephrasal (clue)             | Utilization of a rephrasal due to lack of comprehension or hesitation on the part of the person who is attempting to respond.  |
| Cross-lingual Clue           | Reference to the linguistic features of a language <u>other than</u> the one being learned to help the student find the correct answer.  |
| Translation                  | Elicitation or use of translation to develop the skill or to increase facility in the second language. The <u>ad hoc</u> translation from a language into any other to support further interaction or to increase comprehension. |

*Note.* Fuller explanation of the use of these categories during coding may be found in Gayle (1980a, 1980b).

*Mechanical Use of Language* is characterized by a concern for accuracy, automaticity, and precision; this refers to the type of precision which induces habit formation and which is not necessarily based on a thorough understanding of linguistic structure. It covers the categories complete repetition, partial repetition, and expansion to a full sentence. Experience in second language classrooms has prompted the belief that the combined effect of these categories in the absence of high student aptitudes (such as high ability for language analysis [Pimsleur, 1966]) may merely be to guarantee automatic responses which cannot necessarily promote



transfer of learning to different situations. Research (Gayle, 1982b) has supported the view that students high on the aptitude Language Analysis are better facilitated in aural comprehension by teaching styles which include emphasis on this strategy than are their counterparts low in language analysis.

The fifth strategy, *Direction*, alludes to all teacher behaviours which attempt to manage students by keeping them in touch with the teacher's personal wishes, likes, and dislikes. These behaviours are additional to the information passed on in, for example, corrections. The strategy refers to the categories acceptance, rejection, and directive.

The *Verbal Processing* strategy<sup>3</sup> refers to all categories which demand or give explicit explanations or analyses—that is, categories which help the individual to monitor what he is saying. It covers the categories explicit correction, indication of possible error, clarification, grammar clue, specific information (clue), general information (clue), illustration (clue), rephrasal (clue), cross-lingual clue, and translation. Use of these categories may refer not only to the formal linguistic properties of the previous sentence but also to its meaning.

The merit of discussion based on these strategies rather than on other strategies previously discussed in the literature (e.g., the formal/functional, implicit/explicit, etc. [Stern, 1974]) is that (a) the present strategies are not merely ideational, (b) the discourse functions which contribute to these strategies emerged during empirical classroom observation and were tested throughout the validation of the LTRS for precision to avoid overlap, and (c) with the avoidance of overlap, which would distort the clarity of the configuration and hinder comparisons, the style which emerges is coincident with the total of the discourse functions in their allotted proportion (grouped under strategies) to assume a distinguishable character and a pattern all its own. Style is therefore an instrument which can be measured to explain results in its application to different types of students by different teachers.

The vertical proportions displayed in Table 1 for the teaching styles H, J, and L indicate the relatively stable features of each style and their respective dominance(s). In all styles the *Basic* strategy, characterized by closed-ended questions and answers, has by far the most outstanding percentage frequency. Examination of subtotals in the columns "H" shows that style H is most dominant in the *Basic* strategy and the strategy *Language Exposure* and least outstanding in the strategy *Independent Analysis and Use of Language*. The three remaining intermediate strategies, *Mechanical Use of Language*, *Direction*, and *Verbal Processing*, have disparate strengths among them, *Mechanical Use* having the smallest frequency and *Direction* the largest. Apart from the high frequency for the *Basic* strategy, prominent in all styles, the strongest contributions in style H are from the strategies *Language Exposure* and *Direction*. *Language Exposure* has been defined as extensive experience with the language and observation in second language classrooms suggests that teacher talk is often most spontaneous in the context of the strategy *Direction*. In addition, the strategy *Independent Analysis and Use of Language*, although the weakest numerically in this style, carries a frequency higher than it does in any other style. For these reasons, style H appears to weigh heavily on a certain lack of inhibition or restriction in the use of language and has been assigned the label *OPEN (LATERAL) Language*.

In contrast to style H, once the *Basic* strategy has been accounted for, style L shows greatest dominance in the strategy *Mechanical Use of Language*, least

emphasis still being placed on the strategy *Independent Analysis and Use of Language*. The frequency for *Mechanical Use of Language* in style L is significantly higher than it is for any other style. Moreover, although a frequency of 5.03 is preserved for the strategy *Independent Analysis and Use of Language* in this style, none of this is located in the discourse function, initiation. These facts, together with obvious modulations in frequencies for *Language Exposure* and *Direction* in comparison with style H (Lateral Language) suggest that style L shows relatively stronger tendencies towards restricted use of language and is therefore described as *LINEAR Language*. It should be noted that although this concept seems to rest heavily on the single strategy *Mechanical Use of Language*, this strategy has a frequency almost twice as high as any other strategy in the style.

In style J, the most outstanding strategy (excluding the *Basic* strategy) is *Language Exposure*, the least outstanding is *Independent Analysis and Use of Language*. However, in contrast to style H, the precise numerical proportions show less emphasis on *Language Exposure*, *Direction*, and *Independent Analysis and Use of Language* and more emphasis on *Mechanical Use* and *Verbal Processing*. Since style J retains the use of the discourse function, initiation, which is perhaps the most lateral of all categories and shows frequencies in style J second only to style H, and because style J shows fairly equal proportions on the strategies *Direction*, *Verbal Processing*, and *Mechanical Use* (that is, a mixture of strategies contributing to both lateral and linear language use), style J has been designated *BALANCED*.

The fourth style, K (INFORMATION), which emerged from the cluster analysis, has more relevance to aptitudes (Gayle, 1982b) and, for convenience, has been omitted from Table 1.

A true comparison of styles cannot be made without looking more closely at what the figures in the respective categories disclose on the horizontal plane. When styles H (Lateral Language) and J (Balanced) are compared in terms of horizontal proportions, it is found that style H is preponderant on the strategies *Basic*, *Language Exposure*, *Independent Analysis and Use of Language*, and *Direction*. Predominances in favour of style J on the strategies *Mechanical Use of Language* and *Verbal Processing* are in excess of ten units.

In the H-L comparison (see Table 1) it is found that predominant strategies in style H (Lateral Language) are *Language Exposure*, *Independent Analysis and Use of Language*, and *Direction*. The numerical strength of *Language Exposure* in style H is more than twice that of the same strategy in style L (Linear Language) (a difference of almost 40 units). Most of this dominance must be attributed to the category, reading. Strategies which are horizontally preponderant in style L (Linear Language) are *Basic*, *Mechanical Use of Language*, and *Verbal Processing*. The differential in favour of style L is largest for the strategy, *Mechanical Use of Language* (a difference of just over 40 units). The dominance for the most outstanding horizontal contrasts between styles H and L, is, therefore, almost equally weighted in the two styles.

In the J-L comparison predominant strategies in style J (Balanced) are *Language Exposure*, *Direction*, and *Verbal Processing*. The numerical strength of *Language Exposure* in style J is almost twice that of this strategy in style L and most of this dominance may be claimed by the category, reading. The preponderance in *Language Exposure* is much greater than that in the other two strategies. Preponderance in style L (Linear Language) is seen in the three



strategies *Basic, Independent Analysis and Use of Language*, and *Mechanical Use of Language*.

In no instance do the figures for the strategy *Independent Analysis and Use of Language* exceed 6.64. This was expected since the basic methodology in the Intensive Programme shows strong tendencies towards the audiolingual approach based on the stimulus-response association model. However, as a result of its general rarity and low predictability within the Programme, this strategy could be expected to make a difference wherever it occurs.

Elements related to the discourse functions in styles H, J, and L have been discussed in numerous works which considered the effects of specific teaching techniques or behaviours on learning outcome. For example, the following have been respected—the need for meaningful material (Carroll, 1966), reference to and practice of rules for syntactic structures (Chomsky, 1957; Jakobovits, 1970; Lakoff, 1969; Rivers, 1968), the use of the paraphrase, inference questions, contrasts such as translation, feedback (Fanselow, 1977), writing (Paulston, 1976), and repetition and frequent exposure of basic sounds (Rivers, 1981). In most cases individuals have posited a rationale for the effect of these techniques. Rivers, for example, suggests that foreign language sounds are like a stream of undifferentiated noise. Students are so busy listening for familiar sounds that they may not concentrate on meaning and retention. Repetition and frequent exposure to basic sounds and structures are therefore important elements in teaching for listening comprehension. However, in spite of these and similar postulations regarding the benefits of singular techniques, it remained extremely difficult to make definite predictions regarding the effect of each treatment (teaching style) in this study owing to their complex and diverse character. In order to be consistent with the implications of the concept of teaching style, the differential effects of treatments H, J, and L would depend on the distinct *pattern* of strategies and emphases which was peculiar to each. Since theory on second language teaching style has yet to be articulated, expectations in the form of specific predictions were not entertained.

The problem of relating teaching styles to attitudes towards second language learning was even more tangled. It could be argued that some teaching styles, depending on their emphases in the use of discourse functions and strategies, may confront the low attitude student with more experience outside the orbit of material unique to the French culture. That is, these students may learn to listen for reasons other than French. Speculations are always interesting but often lead astray. They seem even more futile when complex teaching styles are involved. The study was therefore given an exploratory orientation and approached with a minimum of general expectations. The most explicit of these was that since teaching styles demonstrate different configurations of discourse functions used, they would contribute to significant interaction effects with specific attitudes on the criterion variable.

In order to exploit a quasi-experimental design, care was taken to separate the true effects of attitudes and teaching styles on aural comprehension from other potential sources of difference. Reference to the literature on second language learning showed that variables which needed to be measured and controlled because of prior claims of their relevance to language acquisition were motivational intensity (Gardner & Lambert, 1972), second language aptitudes (Carroll, 1966; Pimsleur, 1966) and a general measure of intelligence.



## *Method*

### *Sample*

The original sample of 14 teachers at the grade 6 level in different boards in Ontario volunteered to assist in the study. All teachers participated in the Intensive Programme and offered 75-minute lessons each day per week. Students were instructed in both Social Studies and French Language Arts. Teachers were grouped on the basis of a cluster analysis procedure. Only groups comprised of the same number of teachers (groups H, J, K, and L) were retained for further analysis. Group K is omitted from the present discussion for reasons previously explained. The 118 students who were taught by the nine teachers in groups H, J, and L and who were present for all phases of the testing were used in the analyses described in the following discussion.

### *Instruments*

The Language Teaching Record Scheme (LTRS) (Gayle, 1980a) was used to collect data for the identification of second language teaching styles. The LTRS is a research instrument constructed specifically for the purpose of identifying teaching styles in second language classrooms. Research procedures facilitated by the LTRS differ from those adopted in previous observation of the second language classroom (e.g., Fanselow, 1977; McEwen, 1976) since the LTRS (a) attempts analysis of teaching using categories which are exhaustive yet discrete and (b) simultaneously promotes the study of the effects of combinatory associations among these categories thus utilizing the composite concept referred to as Style.

Both the LTRS and the model on which it was based (Gayle, 1978) were constructed from extensive empirical evidence and classroom data in a number of programmes (inclusive of the Intensive Programme) and boards. The categories of the LTRS are therefore considered to be exhaustive for the type of programmes in which it was validated. The category system it employs refers to a number of dimensions of teaching. The particular dimension of the LTRS (and thus of teaching style) explored in this study is the discourse function which has been defined as the communicative use or role of the message. The discourse function refers only to the *pedagogical* function of the oral behaviour and not to its linguistic function. It therefore specifies teaching techniques used to enhance learning during oral communication and includes student verbal behaviours allowed or encouraged by the teacher. The 24 categories of discourse function used in this study (see Table 2) are not intended to describe the content of the oral communication nor the type of visual aids used, aspects which are described in other dimensions.

Coefficients of observer agreement for the dimension of the LTRS known as the discourse function were .96 (based on Scott's 1955 formula) and .88 (based on Cohen's 1960 formula). This is significant not only in relation to the reliability of the LTRS but also to its validity. The definition of validity offered by Medley and Mitzel (1963) holds that an observation system must provide an accurate record of behaviours which actually occurred in such a way that scores are reliable. The power of the LTRS to record classroom events faithfully is also reflected in the finding (Gayle, 1980b) that among a total of 956 discourse functions selected randomly from two classes only eight ambiguities were discovered; that is, on only

eight occasions did coders see applicability of more than one category on this dimension to the same verbal behaviour in question. Other measures of the instrument's validity include data (Gayle, 1980b) for coefficients of variation across the total validation sample with respect to each category. These data reflect known emphases in the Intensive Programme and the other programmes covered in the validation sample. Perhaps the factor which outweighs all others as an index of the validity of the LTRS is the source of the categories used in the instrument. Categories were not selected arbitrarily from theory nor from those of other observation instruments nor were they conceptualized in abstract. Rather, they emerged as a result of direct observation of classroom events in an extensive validation sample (using over 50 second language classrooms), thus ensuring the relevance of the categories to what actually takes place during teaching. This procedure was followed by inter-observer analysis and evaluation of categories in terms of the existing literature. The content validity of the LTRS is discussed by Gayle (1980b).

A complete description of the empirical procedures used in developing the instrument and a detailed evaluation of its psychometric properties are found in the appendix to the document (Gayle, 1980a) and in the report of a longitudinal study (Gayle, 1980b). Application of the statistical cluster analysis technique to percentage frequencies for all 24 categories of discourse functions allows the identification of groups of teachers who are more homogeneous in their use of discourse functions and thus allows the identification of teaching styles.

Gardner's attitudinal/motivational battery (Gardner & Lambert, 1972) was used to determine student Attitudes towards French Canadians, Attitudes towards Learning French, Desire to Learn French, and Motivational Intensity. The Pimsleur Language Aptitude Battery (Pimsleur, 1966) was used to determine aptitudes for Vocabulary, Language Analysis, Sound Discrimination, and Sound Symbol Association. Subscores on the attitude and aptitude batteries used were treated as individual scores. Scores for general intelligence were derived from the Dominion Group Test of Learning Capacity and measures of the criterion variable, auditory comprehension, were taken from the Quebec Ministry of Education Test of French as a second language. This test was originally constructed for students in the final (sixth) year of elementary school.

### *Procedure*

Teachers in styles H, J, and L (see Table 1) had been identified during the first term of the school year subsequent to three consecutive random visits by trained coders to each of 14 volunteer teachers in the total sample. Coders first transcribed alternate five-minute periods of taped classroom interaction for each class in the total sample. The LTRS was later applied to determine percentage frequencies for categories of discourse functions. Cluster analyses of teachers based on these frequencies identified 4 distinct teaching styles including styles H, J, and L.

Testing of students was not carried out at the beginning of the school year since it was necessary for all teachers to be recruited before simultaneous measurement of students could begin. At the beginning of the Winter Term each class was visited once for the administration of the aural comprehension test (pretest) and Gardner's attitudinal/motivational battery. Tests conducted during three visits to each class in June of the same year were aural comprehension (posttest) and measures of language aptitude and intelligence. Aptitudes and intelligence were not considered

to be sufficiently labile to be significantly affected by prior treatments during the study (Keppel, 1973).

Analyses

As indicated previously, teaching styles H, J, and L were identified through the statistical cluster analysis technique. Data for students were subsequently grouped on the basis of their teacher's teaching style.

Due to the *exploratory* nature of this study and the importance of identifying relationships which may exist among the populations, the decision was made not to apply conventional levels of significance rigidly but to allow a degree of latitude with the intention of increasing the probability of Type I as against Type II errors in the final outcome.

Two-way analyses of covariance were carried out to test for interaction effects between teaching style and attitudes on aural comprehension. The three pairs of teaching styles referred to previously (HJ, HL and JL) were used in turn as the first independent factor. In each case, levels of each of the four attitude/motivation variables identified as Attitudes towards French Canadians, Attitudes towards Learning French, Desire to Learn French, and Motivational Intensity were used in turn as the second independent factor. Cutpoints for these affective variables were identical with the midpoint of the range for each subsample. For each analysis of covariance only those variables which were not being used as an independent factor and which had a linear relationship with the criterion  $\geq .60$  (see Table 3) were actually used as covariates. Adjustments were made for unequal cell sizes in the analyses of covariance.

TABLE 3  
REGRESSION COEFFICIENTS FOR AUDITORY COMPREHENSION  
(POSTTEST) FOR EACH OF THREE SUBSAMPLES  
(PAIRS OF TREATMENT GROUPS)

| Variable                           | HJ    | HL    | JL    |
|------------------------------------|-------|-------|-------|
| Auditory Comprehension (pretest)   | 0.84* | 0.75* | 0.72* |
| Intelligence                       | 0.03  | 0.27* | 0.32* |
| <u>Aptitudes</u>                   |       |       |       |
| Vocabulary                         | 0.22  | 0.09  | 0.29* |
| Language Analysis                  | 0.07  | 0.08  | 0.10  |
| Sound Discrimination               | 0.01  | 0.03  | 0.26* |
| Sound Symbol Association           | 0.16  | 0.29* | 0.30* |
| <u>Attitude/Motivation</u>         |       |       |       |
| Attitudes towards French Canadians | 0.03  | 0.08  | 0.11  |
| Attitudes towards Learning French  | 0.15  | 0.28* | 0.30* |
| Desire to Learn French             | 0.22* | 0.31* | 0.23* |
| Motivational Intensity             | 0.22* | 0.24* | 0.20  |

\*  $p < .05$



### Results and Discussion

As explained previously, absence of theory on second language teaching styles demands present attention to exploratory research rather than the testing of specific hypotheses. One general expectation, however, was that teaching styles such as styles H and L which offered different emphases in the discourse functions used, resulting in distinct behavioural configurations, would interact significantly with attitudes on aural comprehension, the criterion variable.

Only one variable (auditory comprehension pretest) qualified to be used as a covariate with the criterion ( $r \geq .60$ ) (see Table 3).

Extremely large numerical differences between means of the various subgroups were not anticipated since the treatment covered a period of only six months. Differences found were promising and suggest the need for continuing research.

Three of twelve analyses of covariance produced the expected interactions at the .10 level of significance or better (see Table 4). All three interactions were located in designs involving the comparison teaching styles H and L. That is, teaching styles H and L interacted significantly with each of Attitudes towards French Canadians, Attitudes towards Learning French, and Desire to Learn French on aural comprehension. When instructed with teaching style H (Lateral Language), students with relatively low Attitudes towards French Canadians and low Desire to Learn French scored higher on aural comprehension than students with high attitudes. The interactions for these designs were disordinal (see Figure 1 and Table 5). Unlike these two designs, the design involving Attitudes towards Learning French did not show a disordinal interaction; rather, when taught by style H students with relatively low Attitudes towards Learning French did not score quite as high as the high attitude students although they scored appreciably higher than the low attitude students did when taught by style L (Linear Language) (see Figure 1 and Table 5).

Three significant ( $p < .05$ ) main effects were also found. Students with high Attitudes towards Learning French scored consistently higher than low attitude students on aural comprehension in the HL and JL designs. When the factor Desire to Learn French was examined in the JL subgroup, a significant main effect was found for teaching style with style J (Balanced) students scoring higher than style L (Linear Language) students, irrespective of the intensity of their desire to learn the language (see Tables 4 and 5).

Explanation of the tendency for low attitude students to score higher on aural comprehension when taught by teachers in style H (Lateral Language) than when taught by teachers in style L (Linear Language) and, at times, to score even higher than the high attitude students taught by style H, may be offered by the speculation that these students experience style H as an attractive treatment. The possibility exists that the low attitude student's interest is awakened or bolstered by the teacher's relatively greater personal attention to students in style H as may be illustrated in the emphasis placed on a number of strategies. Perhaps students enjoy being called on to read (the strategy *Language Exposure*), to be "on stage." Perhaps they welcome frequent directions and appraisals given by the teacher (the strategy *Direction*). Perhaps they are motivated by the opportunity to take the initiative during the interaction (the strategy *Independent Analysis and Use of Language*). The intense personalized treatment afforded by style H may therefore stimulate students' interest in the spoken word, since they have been inspired to put in their penny's worth as some would say, or to bid for "air time." This would

culminate in higher scores on aural comprehension for low attitude students in spite of their negative attitude. In contrast, style L (Linear Language) would probably engulf students in impersonal choral repetition and drill work. Listening to what is being said would have little pay-off for low attitude students in this latter case.

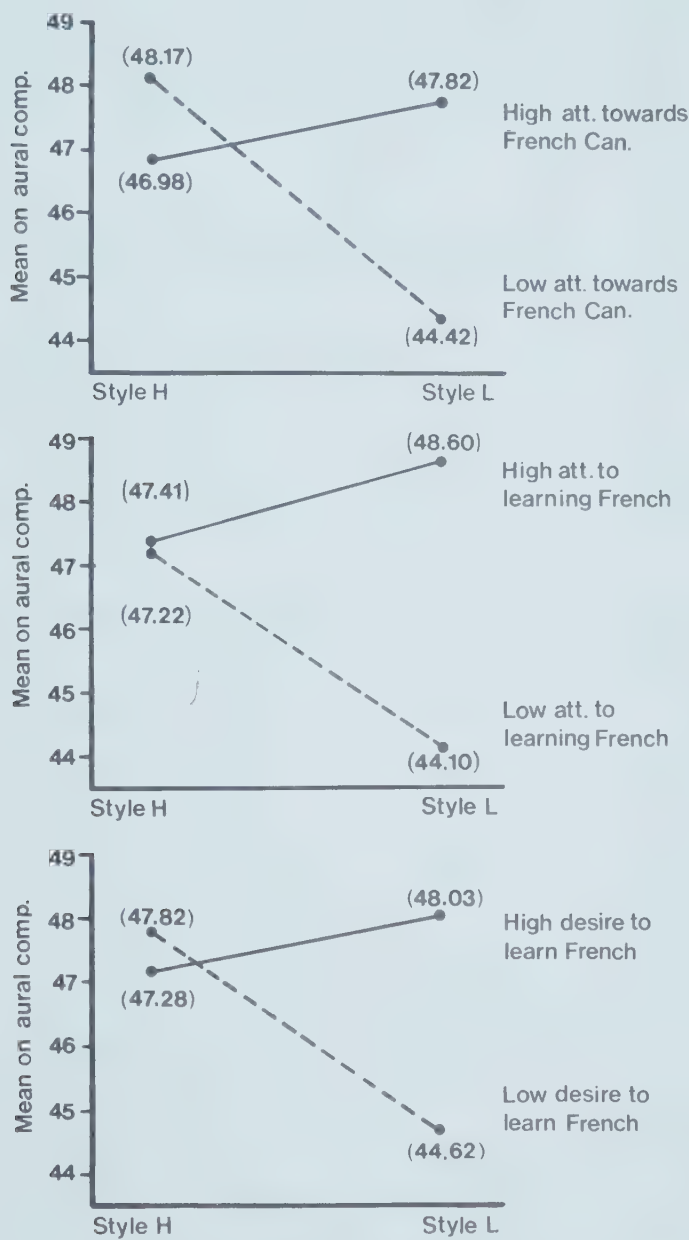


Figure 1. Interactions between teaching style and attitudes on aural comprehension.

A number of possible alternative explanations exist for the higher scores on aural comprehension acquired by low attitude students exposed to style H. The following conjectures may be made:

(a) teaching style affected the second independent variable (attitude) during the actual treatment period. The potential effect of treatment on the second independent variable is discussed in the literature on the aptitude/treatment

interaction (ATI) model (Hunt, 1971; Cronbach & Snow, 1977) and the actual effect of treatment on attitudes is reported in the general research on classroom interaction analysis (Amidon & Flanders 1963; La Shier, 1965; Moskowitz, 1970). Since attitudes are known to have a positive relationship with achievement in the second language (e.g., Burstall, 1975), the effect of teaching style on attitude in this study would have had to occur in the form of a reversal of attitudes (e.g., the low became high and the high became low) in order to have generated the performance scores seen in Figure 1. If this alternative conjecture is accepted, the differential effect of teaching style H is even more apparent than in the first interpretation of the results since style H would be capable of changing attitudes radically.

TABLE 4  
TWO-WAY ANALYSES OF COVARIANCE  
ON AUDITORY COMPREHENSION SCORES

| Source                             | <i>df</i> | <i>MS</i> | <i>F</i> |
|------------------------------------|-----------|-----------|----------|
| Teaching Styles H, L               | 1         | 32.64     | 1.78     |
| Attitudes towards French Canadians | 1         | 21.65     | 1.18     |
| TA                                 | 1         | 92.03     | 5.02 **  |
| Covariate (auditory comp. pretest) | 1         | 1773.19   | 96.71    |
| Error                              | 75        | 18.34     |          |
| Teaching Styles H, L               | 1         | 15.66     | 0.90     |
| Attitudes towards Learning French  | 1         | 93.19     | 5.37 **  |
| TA                                 | 1         | 86.47     | 4.98 **  |
| Covariate (auditory comp. pretest) | 1         | 1306.69   | 75.24    |
| Error                              | 75        | 17.37     |          |
| Teaching Styles H, L               | 1         | 24.35     | 1.33     |
| Desire to Learn French             | 1         | 38.66     | 2.11     |
| TA                                 | 1         | 70.28     | 3.84 *   |
| Covariate (auditory comp. pretest) | 1         | 1818.69   | 99.32    |
| Error                              | 75        | 18.31     |          |
| Teaching Styles J, L               | 1         | 58.21     | 3.30     |
| Attitudes towards Learning French  | 1         | 137.01    | 7.76 **  |
| TA                                 | 1         | 56.99     | 3.23     |
| Covariate (auditory comp. pretest) | 1         | 1180.24   | 66.88    |
| Error                              | 76        | 17.65     |          |
| Teaching Styles J, L               | 1         | 90.54     | 4.80 **  |
| Desire to Learn French             | 1         | 51.06     | 2.71     |
| TA                                 | 1         | 56.47     | 3.00     |
| Covariate (auditory comp. pretest) | 1         | 1774.23   | 94.11    |
| Error                              | 76        | 18.85     |          |

\* Significant at .10 level

\*\* Significant at .05 level or better

(b) the presence of French Canadian teachers in the H group facilitated positive changes in Attitudes towards French Canadians and Desire to Learn French. If this view is being taken, it is pointed out that style L was also comprised exclusively of French Canadian teachers but their presence did not result in the same student performance. It can therefore be safely assumed that performance in style H was not brought about by teachers per se but by the style.



TABLE 5  
DESCRIPTIVE STATISTICS FOR SUBGROUPS  
IN ANALYSES OF COVARIANCE

| Teaching<br>Styles | Subgroups                             | N  | Adjusted<br>Means | SD   |
|--------------------|---------------------------------------|----|-------------------|------|
| H,L                | H, Low Att. towards French Canadians  | 17 | 48.17             | 6.86 |
|                    | H, High Att. towards French Canadians | 19 | 46.94             | 6.16 |
|                    | L, Low Att. towards French Canadians  | 12 | 44.42             | 7.04 |
|                    | L, High Att. towards French Canadians | 32 | 47.82             | 6.14 |
| H,L                | H, Low Att. towards Learning French   | 16 | 47.22             | 6.37 |
|                    | H, High Att. towards Learning French  | 20 | 47.41             | 6.40 |
|                    | L, Low Att. towards Learning French   | 15 | 44.10             | 8.08 |
|                    | L, High Att. towards Learning French  | 29 | 48.60             | 3.32 |
| H,L                | H, Low Desire to Learn French         | 20 | 47.82             | 6.83 |
|                    | H, High Desire to Learn French        | 16 | 47.28             | 5.97 |
|                    | L, Low Desire to Learn French         | 15 | 44.62             | 6.74 |
|                    | L, High Desire to Learn French        | 29 | 48.03             | 6.37 |
| J,L                | J, Low Att. towards Learning French   | 15 | 47.50             | 5.20 |
|                    | J, High Att. towards Learning French  | 22 | 48.64             | 6.61 |
|                    | L, Low Att. towards Learning French   | 15 | 43.84             | 8.08 |
|                    | L, High Att. towards Learning French  | 29 | 48.47             | 3.32 |
| J,L                | J, Low Desire to Learn French         | 19 | 48.54             | 6.67 |
|                    | J, High Desire to Learn French        | 18 | 48.43             | 6.11 |
|                    | L, Low Desire to Learn French         | 15 | 44.41             | 6.74 |
|                    | L, High Desire to Learn French        | 29 | 47.78             | 6.37 |

(c) the view that low attitude students may be responding to the teacher's personality rather than the specific teaching style creates another problem in interpretation of the present results. It is easy to see, however, that teaching style, as defined in this work, is itself a significant part of the teacher's personality, as style reflects basic dispositions towards democratic versus authoritarian attitudes or even priorities for linguistic perfectionism versus contextual inquisitiveness.

It should be noted that high attitude students taught with style H are always somewhat weaker in aural comprehension than high attitude students taught with style L. It is possible that style H is not an incentive for these students. It is equally possible that high attitude students in treatment L (Linear Language) are subject to the illusion, created by the strategy *Mechanical Use of Language* in this style, that they are actually learning how to function in the second language.

It appears, therefore, that while the actual mechanics of the processes which occurred singly or together could only be identified with more complex research, the present research design was sufficiently controlled by the comparison effect to disallow differences in language performance for low versus high attitude students other than those differences which radiated substantially or initially from the styles themselves.

The present data suggest that teaching styles may be capable of breaking the hold of a negative attitude by changing behaviours. Alternatively, they may entirely remodel attitudes. A delicate difference separates the two processes. The important issue is that different styles do produce dissimilar results for specific types of students.

The apparent compatibility between teaching style H and low attitude is not necessarily to be expected for all age groups. Allen and Corder (1974) express the view that the problems of advanced language learners are qualitatively different from those of younger learners. The present writer believes, for example, that it is entirely possible that as individuals grow older, low attitude students in *both* styles H and L could demonstrate a greater tolerance for the strategies of each style without attitudes to French or to French Canadians changing significantly. Conversely, high attitude students could develop a dislike for automatic drill as they grow older.

The fact that only three of twelve analyses of covariance produced the expected interactions at the .10 level of significance or better and that all these interactions related to the HL comparison suggests that, for this age group at least, this particular contrast (Lateral versus Linear Language) is more sensitive to the presence of specific attitudes than other comparisons studied between teaching styles. Further, the consistency noticed with respect to the effect of teaching styles HL on students and their performance is additional evidence of the validity of the LTRS. This view is supported by Rosenshine and Furst (1973) who state that the ultimate test of the validity of an observation instrument rests on whether or not it specifies treatment types which are related to differential student performance.

The main effect on aural comprehension in favour of high positive Attitudes towards Learning French for the HL and JL samples is not particularly surprising in view of previous research (e.g., Gardner & Lambert, 1972) establishing a relationship between attitudes and second language acquisition. A more revealing result was the main effect found for teaching style in the design which used the JL styles and Desire to Learn French as the independent factors.

The effectiveness of teaching style J (Balanced) for aural comprehension irrespective of the strength of the students' Desire to Learn French is reminiscent of the main effect found on aural comprehension in favour of style J (Gayle, 1982b) when styles JL and Aptitude in Sound Symbol Association were the independent factors. Since style J, of the three styles under discussion, has the most equable distribution of frequencies for strategies which are dominant in both Lateral and Linear styles, it could reasonably be expected to produce better results than style L on aural comprehension for all types of students.

These interpretations of the findings in the HL and JL comparisons would be grossly inadequate, however, unless it were added that the effect of any style must be understood in terms of the total configuration of discourse functions within it as well as its particular pattern of dominance over another style. For example, style J (Balanced), when compared with L (Linear Language), represents relatively greater exploitation of the categories 'reading,' 'writing,' and 'initiation' as well as most or all of the categories on the strategies *Verbal Processing* and *Direction*. This unique composition may be regarded as an effective instrument for strengthening aural comprehension.

In spite of the fact that style J is relatively more balanced and gave rise to a main effect in the JL designs, it did not produce a main effect in the HJ comparison. This is probably because the internal pattern of dominance for styles H and J is too similar. It is also worthy of note that the pattern of horizontal predominance for J in the HJ pair is somewhat different from the pattern for J in the JL pair.

This study has supported the view that teaching styles are educational phenomena to be reckoned with not only in the broader educational setting, but also in the narrower second language context. Second Language Teaching Styles, even within the same programme, have been found to make a difference. However, the influence of teaching style does not necessarily stretch across the board. It appears that often the relationship between teaching style and aural comprehension becomes significant when students with special characteristics are considered. Perhaps this is fortunate. Results of this nature do not encourage hasty and erroneous remarks regarding the supposed inherent merit of a particular teaching style over all others. What these results do is to quietly suggest that teaching styles may be sufficiently potent to enhance a good attitude making it even more powerful, or to reduce the already meagre effect of a low attitude to a negative quantity, or to counteract the expected effects of high and low attitudes leading to surprising returns.

It is believed that more than one of these processes may have been implicated in the differential results of this study. However, the reader who wishes to put the information regarding teaching styles under the magnifying glass should do so at the angle which brings the figures under review as a composite whole rather than as a fragmented and disjointed spectacle. In keeping with this principle, it is intended to extend, in the near future, the panorama and to look simultaneously at other dimensions of teaching, including the potency of interactivity among components in a style.

When the question of Style is discussed as an integrated composite the uniqueness of teachers as Artists becomes apparent. There is now room for creativity and their contribution to the learning process is etched in bolder outline, making them distinguishable above fellows of the same School who use the same brush, the same colours, and the same palette. It is not surprising that the dominant theme in a Teaching Style can be portrayed and can be identified with student production. Nor should it be surprising that, whereas the mortal may be unable to “dip his brush in dyes of heaven” as Scott says in his *Marmion*, yet there are fetching pastel shades and alluring tones available to the teacher which can be interspersed to catch the imagination of the student. The subject, therefore, of Styles defies complete orthodoxy. Teachers must first analyze their own style to find what it really achieves, must find the style best suited for their gifts, and must alter it if necessary so that Style may be the effective tool it can be in making or breaking attitudes.

#### Notes

1. It is to be understood that the two Rs (reading and writing) have necessarily to be recognized and included in discourse functions where they are interspersed with oral interaction.
2. All teachers in the present study were taken from this Programme. The Intensive Programme in French as a second language is organized to accommodate both Language Arts and Social Studies. This allows increased opportunity for the actual use of the second language which is the sole medium of communication during classroom interaction.
3. The *Verbal Processing* strategy differs from the Formal strategy (Stern, 1974) since it not only helps students study the language and monitor their use of the linguistic properties of the sentence, but it may also induce reflection on the content of what is being said.

This research was funded by the Faculty of Education, University of Ottawa.



References

- Allen, J. P. B., & Corder, S. P. (1974). *Techniques in applied linguistics. Edinburgh course in applied linguistics* (Vol. 3). London: Oxford University Press.
- Amidon, E. J., & Flanders, N. A. (1963). *Role of the teacher in the classroom*. Minneapolis: Paul Amidon Associates.
- Burstall, C. (1975). French in the primary school: The British experiment. *Canadian Modern Language Review*, 31, 388-402.
- Canfield, R. (1978). *Instructional style inventory*. Ann Arbor: Humanics Media.
- Carroll, J. B. (1966). The contributions of psychological theory and educational research to the teaching of foreign languages. In A. Valdman (Ed.), *Trends in language teaching*. New York: McGraw-Hill.
- Carroll, J. B., & Sapon, S. M. (1967). *Modern Language Aptitude Test (MLAT)*. New York: The Psychological Corporation.
- Chastain, K. (1976). *Developing second language skills*. Toronto: Rand McNally.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Cohen, J. C. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671-684.
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods*. New York: Irvington.
- Fanselow, J. F. (1977). Beyond RASHOMON—Conceptualizing and describing the teaching act. *TESOL Quarterly*, 11, 17-39.
- Fanselow, J. F., & Light, R. (1977). *Bilingual, ESOL and foreign language teacher preparation: Models, practices, issues*. Washington: TESOL.
- Flanders, N. A. (1970). *Analyzing teaching behaviour*. Reading, Mass.: Addison-Wesley.
- Gardner, R. C., Gliksman, L., & Smythe, P. C. (1978). Attitudes and behaviour in second language acquisition: A social psychological interpretation. *Canadian Psychological Review*, 19, 173-186.
- Gardner, R. C. & Lambert, W. E. (1972). *Attitudes and motivation in second language learning*. Rowley: Newbury House.
- Gayle, G. M. H. (1978). A revised model of second language teaching. Presented at the Learned Societies Conferences, University of Western Ontario, May 21 - June 3.
- Gayle, G. M. H. (1980a). *The language teaching record scheme (LTRS)—Pedagogical functions*. Research instrument, University of Ottawa.
- Gayle, G. M. H. (1980b). *A descriptive analysis of second language teaching styles in the oral approach*. Ottawa: University of Ottawa Press.
- Gayle, G. M. H. (1982a). Identifying second language teaching styles: A new approach to an old problem. *Canadian Modern Language Review*, 38, 254-261.
- Gayle, G. M. H. (1982b). Second language teaching styles and student aptitudes. Research paper, University of Ottawa.
- Hunt, D. E. (1971). *Matching models in education*. Toronto: OISE.
- Jakobovits, L. A. (1970). *Foreign language learning: A psycholinguistic analysis of the issues*. Rowley, Mass.: Newbury House.
- Katz, D., & Stotland, E. A. (1959). A preliminary statement to a theory of attitude structure and change. In S. Koch (Ed.), *Psychology, a study of a science* (Vol. 3). New York: McGraw-Hill.
- Keppel, G. (1973). *Design and analysis: A researcher's handbook*. New Jersey: Prentice-Hall.

- Lakoff, R. (1969). Transformational grammar and language teaching. *Language Learning*, 19, 117-140.
- La Shier, W. S. (1965). An analysis of certain aspects of the verbal behaviour of student teachers of eighth grade students participating in a BSCS Laboratory Block. Ph.D. Dissertation, University of Texas.
- Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw-Hill.
- Mackey, W. F. (1965). *Language teaching analysis*. London: Harlow, Longmans, Green.
- Mackey, W. F. (1967). Method analysis. A survey of its development, principles and techniques. In J. E. Alatis (Ed.), *Georgetown University Monograph No. 18 Languages and Linguistics*. Washington, DC: Georgetown University Press.
- McEwen, N. (1976). An exploratory study of the multidimensional nature of teacher-student verbal interaction in second language classrooms. Ph.D. Dissertation, University of Alberta.
- Medley, D. M., & Mitzel, H. E. (1963). Measuring classroom behaviour by systematic observation. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally.
- Moskowitz, G. (1970). A comparison of foreign language student teachers trained and not trained in interaction analysis. *American Foreign Language Teacher*, 1, 10-15.
- Moskowitz, G. (1976). The classroom interaction of outstanding foreign language teachers. *Foreign Language Annals*, 9, 135-157.
- Mueller, T. H., & Miller, R. (1970). A study of student attitudes and motivation in a collegiate French course using programmed language instruction. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 8, 297-320.
- Myers, I. B. (1980). *Gifts differing*. Palo Alto, California: Consulting Psychologists Press Inc.
- Orme, M. *Effective teaching techniques*. Videotape series. Toronto: Ryerson Polytechnical Institute.
- Orme, M. (1978). *Teaching Strategies Kit*. Toronto: Department of Applied Psychology, OISE.
- Paulston, C. B. (1976). *Teaching English as a second language: Techniques and procedures*. Cambridge, Mass: Winthrop.
- Pervin, L. A. (1968). Performance and satisfaction as a function of individual-environment fit. *Psychological Bulletin*, 69, 56-68.
- Pimsleur, P. (1966). *Pimsleur Language Aptitude Battery, Form S*. New York: Harcourt, Brace, Jovanovich.
- Rivers, W. M. (1968). *Teaching foreign language skills*. Chicago: University of Chicago Press.
- Rivers, W. M. (1981). *Teaching foreign language skills* (2nd ed.). Chicago: University of Chicago Press.
- Rosenshine, B., & Furst, N. (1973). The use of direct observation to study teaching. In R. M. W. Travers (Ed.), *Second handbook of research on teaching*. Chicago: Rand McNally.
- Scott, W. A. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly*, 19, 321-325.
- Stern, H. H. (1974). Directions in language teaching theory and research. In J. Quistgaard, H. Schwarz & H. Spang-Hanssen (Eds.), *Applied linguistics problems and solutions*. Proceedings of the Third International Congress of Applied Linguistics (Vol. III). Heidelberg: Julius Groos.
- Tucker, G. R., Hamayan, E., & Genesee, F. H. (1975). Affective, cognitive and social factors in second language acquisition. Research paper, McGill University.

HORST H. MUELLER

*The University of Alberta*

UDAYA N. DASH

*Utkal University, India*

DAVID W. MATHESON

and

ROBERT H. SHORT

*The University of Alberta*

## WISC-R Subtest Patterning of Below Average, Average, and Above Average IQ Children: A Meta-Analysis

*Research on WISC-R subtest performance patterns of both retarded and above average IQ children has revealed a number of distinct differences. Retarded children perform relatively better on the Performance Scale subtests than on the Verbal Scale subtests while above average IQ children demonstrate a greater facility for Verbal over Performance Scale subtests. Although the research with exceptional groups appears well substantiated, few, if any, studies directly compare WISC-R profiles of exceptional children with those of normal children of average ability.*

*This study used meta-analytical procedures to compare WISC-R subtest performance patterns of 29 samples of retarded, average, and above average IQ children drawn from the recent research literature. The analysis resulted in three distinct WISC-R subtest profiles for the three groups. The retarded and above average children performed in line with previous research findings while normal children revealed a*

---

Horst Mueller (MA, University of Saskatchewan) is presently a doctoral student in Special Education within the Department of Educational Psychology at the University of Alberta. His research interests include behavior therapy with children, self-management training and behavioral medicine, clinical assessment, and reading comprehension processes.

Dr. Udaya Dash is presently a lecturer in Psychology at Utkal University, Bhubaneswar, India. His major research interests are in the areas of cross-cultural psychology and the effects of schooling, individual differences, and psychometrics.

David Matheson (BA, University of Saskatchewan) is currently completing requirements for an MEd in Educational Psychology at the University of Alberta. His research interests include cross-cultural psychology, cognition, and research design.

Dr. Robert Short is Associate Professor within the Department of Educational Psychology at the University of Alberta. His research interests include human learning and development, individual differences, and instructional psychology.



*subtest score pattern reflecting aspects of both retarded and above average IQ children.*

*The study demonstrated that WISC-R subtest performance patterns are strongly related to overall intellectual level. The comparative performance of children across the individual WISC-R subtests and across subtest clusterings and factor groupings appears to change as a function of Full Scale IQ.*

The *Wechsler Intelligence Scale for Children- Revised* (WISC-R) is presently the most popular psychometric instrument for the assessment of intellectual ability in individual children—both normal and retarded (Kaufman, 1980; Quattrocchi & Sherrets, 1980). Although the WISC-R standardization sample (Wechsler, 1974) was limited to “normal” children only and included no institutionalized or severely disturbed children, a number of studies have shown the reliabilities and stability of the three WISC-R Scales and most WISC-R subtest scores to be satisfactory for use with mildly and moderately retarded children (Covin, 1977; Lutey, 1977; Sattler, 1982; Vance, Gaynor & Coleman, 1977). Research has also shown the WISC-R to have essentially the same factor structure for retarded children (Van Hagen & Kaufman, 1975; Schooler, Beebe & Koepke, 1978) as for normal (Kaufman, 1975, 1980; Wallbrown, Blaha, Wallbrown & Engin, 1975) and gifted children (Karnes & Brown, 1980), thus supporting the notion of a relatively invariant structure of intelligence for school-age children regardless of their overall level of intellectual ability. Further, this WISC-R factor structure has been shown to be relatively stable across sex, socioeconomic status, age, and clinical diagnoses (Kaufman, 1979a, 1979b, 1980, 1981; Lutey, 1977; Quattrocchi & Sherrets, 1980; Sattler, 1982). At a practical level, such findings suggest that children from a broad range of educational and socioeconomic backgrounds, demonstrating a wide variety of academic and interpersonal behaviors, perceive the stimulus materials and tasks of the WISC-R in a similar fashion (Schooler et al., 1978).

However, this is not to say that there are absolutely no differences in the WISC-R performance of such exceptional groups as the mentally retarded or the gifted when compared to “normal” groups. Regardless of the similarity of the WISC-R factor structure across different groups of children, there is ample evidence that various exceptional groups do demonstrate differing patterns of performance (high vs. low scores) across the individual subtests of the WISC-R, and such characteristic patterns of subtest performance are viewed by many researchers as indicative of specific mental operations or brain-behavior relationships (Kaufman, 1980; Sattler, 1982). Thus, in spite of the fact that the WISC-R was designed solely as a measure of general intelligence (Wechsler, 1974), researchers have spent much time and energy searching for subtest performance profiles characteristic of specific clinical categories such as: the mentally retarded (e.g., Kaufman & Van Hagen, 1977; Naglieri, 1980; Vance, Hankins, Wallbrown, Engin & McGee, 1978), the learning disabled (e.g., Anderson, Kaufman & Kaufman, 1976; Moore & Wielen, 1981; Vance, Gaynor & Coleman, 1976), the emotionally disturbed (e.g., Dean, 1977, 1978; Hamm & Evans, 1978; Morris, Evans & Pearson, 1978), the delinquent (e.g., Hays & Solway, 1977; Hays, Solway & Schreiner, 1978; Ollendick, 1979), the intellectually superior (e.g., Karnes & Brown, 1980; Schiff, Kaufman & Kaufman, 1981; Wheaton & Vandergriff, 1978), and the neurologically impaired (e.g., Kuncie & McMahon, 1979; Simensen & Sutherland, 1974; Swerdlik & Wilson, 1979; Tew, 1977) and, over the years, much interest has been generated in

applications of the WISC-R as a diagnostic aid for the classification of exceptional children (Kaufman, 1979a; Mueller, Mancini & Short, 1983; Quattrocchi & Sherrets, 1980; Vance & Singer, 1979). The results of such studies have been less than totally consistent (Dudley-Marling & Kaufman, 1981; Lutey, 1977; Sattler, 1982) and some controversy has arisen in regard to the validity of the WISC-R profile analysis approach to the classification of individuals (Berk, 1983; Gutkin, 1979; Hale & Saxe, 1983; Hirshoren & Kavale, 1976; Miller, 1980; Miller & Walker, 1981; Ryckman, 1981). This inconsistency largely exists because, although researchers have noted statistically significant WISC-R profiles that appear characteristic of certain exceptional groups, they have failed to demonstrate these profiles to be distinctive enough to differentiate between exceptional and normal children (Hale & Saxe, 1983). Researchers, such as Berk (1983), further argue that the application of evidence for characteristic group-profiles to the individual child is both problematic and fraught with sources of error. However, other researchers continue to see potential in further research within this area, arguing that a better understanding of the processes underlying differential performance on the WISC-R subtests may lead to more comprehensive understanding of the specific problems, strengths and weaknesses of various exceptional groups (Kaufman, 1979a, 1980, 1981; Sattler, 1982).

In regard to the mentally retarded, the literature suggests uneven patterns of performance on tests of intelligence when compared to children of average intellect. Early studies of the WISC performance of retarded children noted that Vocabulary, Arithmetic, and Information were consistently among the lowest subtest scores for these children, whereas Object Assembly and Picture Completion were typically highest (Baumeister, 1964; Silverstein, 1968; Thompson & Finley, 1962). More recent studies using the WISC-R (Kaufman & Van Hagen, 1977; Naglieri, 1980; Vance et al., 1978), while not in total agreement on the relative rankings of all the WISC-R subtests, have generally found a pattern similar to that reported for retarded children on the WISC. Kaufman and Van Hagen (1977) found three Performance Scale subtests—Picture Completion, Object Assembly, and Block Design—to be easiest for the children in their sample; with Vocabulary and Arithmetic the most difficult. Naglieri (1980) also found Object Assembly to be easiest and Arithmetic to be the most difficult for his retarded sample. Similarly, Lutey (1977), reporting the WISC-R subtest consensus ranking across 7 groups of retarded children, found Object Assembly and Picture Completion to be the highest subtest scores in 6 of the groups and Vocabulary and Information to be the lowest scores in 4 and 5 groups, respectively. On the other hand, Vance and his colleagues (1978) reported Picture Completion and Object Assembly to be the second and third easiest, respectively, and Information to be the most difficult subtest for their sample. Thus, ignoring some inconsistency in the exact ordering of WISC-R subtests from one sample of retarded children to another, it does appear that retarded children as a group generally do less well on the Verbal Scale subtests than on the Performance Scale subtests (Lutey, 1977; Vance et al., 1978).

A number of researchers (e.g., Clarizio & Bernard, 1981; Henry & Wittman, 1981; Smith, Coleman, Doeckci & Davis, 1977a, 1977b; Webster & Lafayette, 1980) have also attempted to determine whether mentally handicapped children show characteristic WISC-R performance patterns when subtests are grouped according to Bannatyne's (1974) recategorization system. Such studies have generally found retarded children to score highest on the *Spatial* cluster (Object Assembly, Block Design, Picture Completion) and lowest on the *Acquired Knowledge* cluster



(Information, Arithmetic, Similarities). There has been less consistency regarding the rank ordering of the *Conceptual* (Similarities, Vocabulary, Comprehension) and the *Sequential* (Arithmetic, Digit Span, Coding) clusters.

In contrast to the WISC-R pattern demonstrated by the retarded, the results of studies with children of above average intelligence (Karnes & Brown, 1980; Schiff et al, 1981; Thompson & Finley, 1962; Wheaton & Vandergriff, 1978) suggest that bright children tend to perform best on the Verbal Scale subtests and less well on the Performance Scale subtests—a pattern almost completely opposite that shown by their retarded counterparts. Again, the exact rank-ordering of the subtests varies from sample to sample but, clearly, the intellectually above average show relative strength on the Verbal Scale subtests that together define Kaufman's (1975) *Verbal Comprehension* factor—Information, Similarities, Vocabulary, and Comprehension. Similarly, Schiff, Kaufman and Kaufman (1981) reported that gifted learning disabled students scored highest on Bannatyne's *Conceptual* cluster and lowest on *Sequential*. The only point of similarity between the WISC-R subtest patterns of the retarded and the intellectually superior is that both groups seem to find the Arithmetic subtest to be relatively difficult.

The differences in the apparent patterning of WISC-R subtest scores in these two groups has been explained in various ways. Silverstein (1968) suggested that these differences may be due to the relatively concrete thinking patterns associated with mental retardation, whereas Alper (1967) suggested that retarded children may do more poorly on the Verbal subtests because of inadequate language skills and schooling experiences. Relatively poorer performance on the WISC-R Performance subtests by the intellectually bright may be explained by Kaufman's (1979c) finding that speed of performance becomes an increasingly important factor for high scores on many of the Performance subtests as children become older or for younger gifted children. At most age levels, superior scale scores on Block Design, Object Assembly, Picture Arrangement, and Coding can only be obtained by receiving some of the bonus points available for rapid test performance.

Although it appears that retarded children demonstrate a different pattern of WISC-R performance than do the intellectually superior, it is not yet clear how the patterns of the retarded and superior differ from that shown by normal children of average ability. Kaufman (1976c) presented data showing that the normal population demonstrates a surprising amount of subtest scatter on the WISC-R. Apparently, it is not particularly unusual for as many as four WISC-R subtests to significantly vary from the mean scaled score (based on 12 subtests) of a child in the normal population (Kaufman, 1976a). Since normal children do not demonstrate a completely flat WISC-R profile, the question remaining is whether or not there is a characteristic pattern of subtest performance for these children and whether such a pattern more closely resembles the pattern of retarded children or superior children. Another question to be answered is how WISC-R subtest rankings change with increasing general intelligence—from retarded to average to superior.

The present study uses meta-analytical procedures (Glass, 1976) to examine the WISC-R subtest performance patterns of 29 samples of below average, average, and above average IQ children drawn from the research literature on the WISC-R to determine whether there exist reliable subtest performance profiles characteristic of different levels of intellectual ability. Meta-analysis refers to the analysis of analyses or, in other words, to the statistical analysis of a large collection of results



from individual studies for the purpose of integrating findings. By comparing and contrasting a number of samples representing different intellectual classifications, the degree to which profiles uniquely and consistently differ, as well as the similarities among subgroups, may be ascertained.

TABLE 1  
SOURCE AND DESCRIPTION OF THE 29 SAMPLES  
SELECTED FOR META-ANALYSIS

| Sample and Source               | Mean<br>FSIQ | Sample<br>Size | Description   |
|---------------------------------|--------------|----------------|---|
| 1 Henry & Wittman (1981)        | 63           | 40             | EMH, mixed race and sex                             |
| 2 Schooler et al. (1978)        | 69           | 127            | EMH, mixed race & sex                               |
| 3 Clarizio & Bernard (1981)     | 69           | 141            | EMH, mixed race & sex                               |
| 4 Rowe (1977)                   | 60           | 44             | EMH, Australian, mixed sex                          |
| 5 Gironde (1977)                | 64           | 20             | EMH, mixed race & sex                               |
| 6 Vance et al. (1978)           | 62           | 238            | EMH, mixed race & sex                               |
| 7 Catron & Catron (1977)        | 60           | 29             | EMH, mixed race & sex                               |
| 8 Kaufman & Van Hagen (1977)    | 51           | 80             | EMH, mixed race & sex                               |
| 9 Vance (1979)                  | 63           | 126            | EMH, wtd. mean of males & females, mixed race       |
| 10 Nagle & Lazarus (1979)       | 63           | 30             | EMH, 16 yrs. old, mixed race & sex                  |
| 11 Mueller (note 2)             | 62           | 20             | EMH, mixed race & sex                               |
| 12 Morris, et al. (1980)        | 63           | 78             | EMH, mixed race & sex                               |
| 13 Raviv et al. (1981)          | 101          | 60             | Normals, Israeli boys                               |
| 14 Hale & Landino (1981)        | NA           | 34             | Unimpaired clinic., mixed race, males only          |
| 15 Pristo (1978)                | 98           | 40             | Normals, mixed race & sex                           |
| 16 Schwarting (1976)            | 106          | 58             | Normals, mixed race & sex                           |
| 17 Stokes et al. (1978)         | 107          | 59             | Normals, gr.6, mixed race & sex                     |
| 18 Schooler et al. (1978)       | 92           | 269            | Unimpaired clinic., mixed race & sex                |
| 19 Clarizio & Bernard (1981)    | 93           | 294            | Unimpaired clinic., mixed race & sex                |
| 20 Peters (1976)                | 107          | 300            | Normals, Canadian, wtd. mean of 3 groups, mixed sex |
| 21 Sandoval (1982)              | NA           | 332            | Normals, Caucasian, mixed sex                       |
| 22 Gutkin & Reynolds (1981)     | NA           | 1868           | Normals, Caucasian, mixed sex                       |
| 23 Resnick (1977)               | 97           | 76             | Normals, mixed race & sex                           |
| 24 Wersh & Briere (1981)        | 116          | 91             | "Normal", Canadian, mixed sex                       |
| 25 Mueller (note 2)             | 121          | 28             | Unimpaired high IQ clinic., Caucasian, mixed sex    |
| 26 Schiff et al. (1981)         | 123          | 30             | Gifted LD, mixed race & sex                         |
| 27 Wheaton & Vandergriff (1978) | 140          | 26             | Gifted, mixed race & sex                            |
| 28 Karnes & Brown (1981)        | 126          | 946            | Gifted, mixed race & sex                            |
| 29 Henry & Wittman (1981)       | 130          | 40             | Gifted, mixed race & sex                            |

#### Method

Standard literature search procedures were utilized to locate research studies in which the WISC-R was administered to samples of children designated as either subnormal or mentally retarded, or average or normal or unimpaired, or above average or superior or gifted. Samples of children described as being predominantly brain-damaged, emotionally disturbed, learning disabled, or of ethnic minorities were excluded from analysis. The one exception, the Schiff et al. (1981) sample of

gifted learning disabled children, was included because of the dearth of “normal” gifted samples in the research literature. Subsamples of children designated as unimpaired or nonhandicapped drawn from larger clinic-referred samples were also accepted for inclusion within the present analysis. All selected samples were required to have a sample-mean FSIQ below 70, between 90 and 109, or above 115. A total of 12 below average, 11 average, and 6 above average IQ samples comprising a total of 5574 children were found suitable for use in this study. These 29 samples are referenced within Table 1.

For each of the 29 selected samples, the sample-mean scaled scores for each of the 10 regular WISC-R subtests (Digit Span & Mazes excluded) were used as the data-base (see Table 2).

TABLE 2  
RAW SCALED SCORE DATA FOR THE 29 SAMPLES ANALYZED

| Sample | Verbal Tests |      |      |      |      | Performance Tests |      |      |      |      |
|--------|--------------|------|------|------|------|-------------------|------|------|------|------|
|        | Inf.         | Sim. | Ar.  | Voc. | Com. | P.C.              | P.A. | B.D. | O.A. | Cod. |
| 1      | 4.0          | 4.6  | 4.0  | 4.6  | 5.2  | 6.2               | 3.8  | 4.2  | 5.7  | 4.4  |
| 2      | 4.7          | 4.6  | 4.9  | 4.7  | 5.4  | 6.9               | 5.7  | 5.4  | 6.4  | 5.1  |
| 3      | 4.7          | 4.7  | 4.9  | 4.4  | 5.3  | 6.9               | 5.7  | 5.6  | 6.2  | 5.1  |
| 4      | 4.0          | 3.2  | 3.3  | 3.5  | 4.8  | 6.7               | 5.1  | 5.2  | 5.4  | 3.4  |
| 5      | 2.7          | 4.0  | 5.1  | 4.6  | 4.9  | 5.2               | 5.1  | 4.2  | 6.5  | 4.7  |
| 6      | 3.5          | 3.6  | 4.8  | 4.0  | 5.5  | 5.4               | 4.0  | 4.2  | 5.0  | 4.3  |
| 7      | 3.2          | 3.3  | 3.8  | 3.6  | 5.0  | 5.2               | 4.3  | 3.8  | 5.0  | 4.3  |
| 8      | 2.8          | 3.0  | 2.1  | 2.0  | 3.0  | 3.8               | 2.7  | 3.2  | 3.2  | 2.6  |
| 9      | 3.6          | 3.6  | 4.9  | 3.9  | 5.5  | 5.6               | 4.0  | 4.3  | 4.9  | 4.3  |
| 10     | 3.8          | 4.3  | 4.1  | 2.9  | 3.9  | 5.6               | 5.2  | 4.9  | 6.5  | 5.2  |
| 11     | 3.3          | 3.1  | 3.7  | 3.3  | 4.0  | 6.9               | 3.3  | 5.2  | 5.8  | 3.9  |
| 12     | 3.6          | 3.5  | 5.1  | 4.1  | 5.2  | 6.3               | 3.7  | 4.2  | 5.3  | 5.3  |
| 13     | 9.9          | 9.8  | 10.2 | 10.0 | 10.5 | 10.1              | 10.3 | 9.7  | 10.7 | 10.0 |
| 14     | 8.9          | 8.8  | 8.2  | 9.3  | 9.2  | 10.1              | 11.3 | 9.7  | 10.7 | 6.5  |
| 15     | 9.0          | 9.6  | 9.8  | 9.3  | 8.4  | 10.0              | 10.4 | 9.6  | 11.6 | 8.6  |
| 16     | 10.2         | 10.5 | 10.1 | 11.6 | 11.7 | 10.5              | 11.2 | 10.7 | 11.5 | 10.6 |
| 17     | 10.6         | 11.3 | 11.4 | 11.1 | 10.9 | 11.0              | 11.3 | 11.2 | 11.3 | 10.5 |
| 18     | 8.1          | 8.2  | 8.2  | 8.6  | 9.1  | 9.9               | 9.5  | 9.0  | 9.7  | 8.2  |
| 19     | 8.1          | 8.3  | 8.2  | 8.7  | 9.1  | 9.9               | 9.5  | 9.1  | 9.6  | 8.3  |
| 20     | 10.4         | 10.7 | 11.0 | 11.0 | 11.3 | 11.1              | 11.3 | 11.1 | 11.0 | 10.9 |
| 21     | 10.4         | 10.3 | 10.2 | 10.8 | 10.0 | 10.7              | 10.3 | 11.1 | 10.5 | 10.7 |
| 22     | 10.4         | 10.3 | 10.4 | 10.4 | 10.4 | 10.4              | 10.4 | 10.4 | 10.4 | 10.2 |
| 23     | 9.1          | 10.3 | 9.0  | 10.8 | 9.4  | 9.7               | 9.6  | 9.5  | 10.4 | 9.1  |
| 24     | 11.9         | 12.5 | 11.3 | 12.2 | 12.6 | 12.5              | 12.6 | 12.2 | 12.6 | 11.1 |
| 25     | 12.3         | 14.1 | 12.3 | 12.6 | 13.5 | 12.7              | 12.6 | 13.1 | 13.0 | 11.4 |
| 26     | 14.6         | 16.2 | 12.1 | 15.5 | 14.4 | 12.2              | 13.0 | 12.7 | 12.4 | 9.6  |
| 27     | 16.1         | 16.9 | 14.4 | 15.7 | 17.1 | 14.7              | 14.9 | 15.5 | 15.7 | 12.8 |
| 28     | 14.0         | 14.5 | 13.4 | 14.7 | 14.5 | 13.2              | 13.3 | 13.0 | 13.2 | 12.2 |
| 29     | 14.2         | 14.8 | 13.1 | 15.5 | 15.7 | 13.4              | 13.5 | 13.9 | 13.9 | 11.8 |

The analysis of the scaled score data proceeded through a number of discrete steps. First, the set of 29 by 10 scaled scores was subjected to *K*-means clustering by way of the BMDPKM program (Dixon, 1981). The PKM procedure partitions a set of *n* cases or samples, each consisting of *k* variable scores, into clusters. At the completion of the procedure each sample belongs to the cluster whose center is closest to the sample. The Euclidean distance (*D*<sup>2</sup>) (Cronbach & Gleser, 1953) is

used to measure the distance between the sample and the center of each cluster. The PKM procedure begins with all the samples in one cluster and splits one cluster into two clusters at each step. Clustering continues until a user-specified number of clusters is reached, at which point, samples are iteratively reallocated into the cluster whose center is closest. Analysis was performed for 2, 3, 4, 5, and 6-cluster solutions. The solution resulting in the smallest number of clusters (minimum  $n=2$  samples) of maximum homogeneity while simultaneously maximizing the average distance between clusters was accepted as optimal.

Groups of samples formed empirically by the clustering procedure were then subjected to a One-Way Multivariate Analysis of Variance (MANOVA) procedure (see Finn, 1974; Morrison, 1967). The program used (Hunka, 1980) carries out a one-way MANOVA on  $n$  groups by  $k$  dependent variables and permits the user to input contrast matrices for special contrasts and/or form linear combinations of dependent variables. Confidence intervals for each pair of group differences on each variable mean (including user-created variables made up of linear combinations of input variables) are calculated. Tests include Rao's approximate  $F$  and Heck's test. Main effects tests include a Helmert contrast matrix which is evaluated row by row. The MANOVA test for the equality of the vector of the variable means was followed by the production of simultaneous 95% confidence intervals for pairwise comparisons on each variable. A MANOVA main effects test of the parallelism of the profile line-segments across groups was permitted by the creation of 9 linear combinations of variables for each of the three groups. Each combination consisted of variable  $k$  minus variable  $k+1$  (e.g., Information-Similarities, Similarities-Arithmetic, Arithmetic-Vocabulary, etc.). Each combination was tested for equality across groups and simultaneous confidence intervals contrasting pairs of groups on each profile line-segment at  $\alpha .05$  were output.

### Results

Table 1 presents the source and description of the 29 samples selected for the present meta-analysis. As may be seen, the sample-mean FSIQs range from a low of 51 to a high of 140 and the selected samples represent a fairly heterogeneous mix of children through this range of IQ scores.

Table 3 presents the results of the  $K$ -means clustering procedure for 2, 3, 4, and 5-cluster solutions. An examination of the triangular matrices of sample-mean Euclidean distances ( $D^2$ ) within clusters versus the mean Euclidean distances between clusters across the presented clustering solutions suggests the 3-cluster solution to be the most optimal. The 5-cluster solution adds little to increased homogeneity of the clusters formed and has the disadvantage of creating a separate cluster consisting of only one sample, while the 2-cluster solution demonstrates inadequate relative homogeneity for the first cluster. The 4-cluster solution does appear to increase cluster homogeneity to some extent over the 3-cluster solution, but there is relatively poor separation between the second and third clusters. However, the 4-cluster solution may also be considered adequate.

Basically, the 3-cluster solution groups all 12 of the below average IQ samples into one cluster, all 11 of the average IQ samples into another, and all 6 of the above average IQ samples into a third cluster on the basis of similarity between sample WISC-R subtest profiles. The 4-cluster solution maintains the same below average and above average groupings of samples but splits the average IQ samples



into two separate groups: a lower average group and a higher average group. The subtest means, grand mean of the subtests, and standard deviation of the subtest means for each of the sample-groups defined by the 3- and 4-cluster solutions are presented in Table 4.

TABLE 3  
CLUSTERING DATA FOR TWO, THREE, FOUR  
AND FIVE-CLUSTER SOLUTIONS

| Cluster | Samples<br>in<br>Cluster | Cluster<br>Composition | Mean Distances in<br>Standard Units<br>Between and<br>Within Clusters |      |      |      |     | Range of<br>Distances<br>from Cluster<br>Center |
|---------|--------------------------|------------------------|---|------|------|------|-----|---|
|         |                          |                        | 1   | 2    | 3    | 4    | 5   |   |
| 1       | 17                       | 11B+6C                 | 1.44  | --   | --   | --   | --  | .36-3.49  |
| 2       | 12                       | 12A                    | 5.69  | .61  | --   | --   | --  | .35-1.54  |
| 1       | 6                        | 6C                     | .87   | --   | --   | --   | --  | .34-1.61  |
| 2       | 11                       | 11B                    | 2.95  | .74  | --   | --   | --  | .31-1.11  |
| 3       | 12                       | 12A                    | 7.58  | 4.67 | .61  | --   | --  | .35-1.54  |
| 1       | 6                        | 6C                     | .87   | --   | --   | --   | --  | .33-1.61  |
| 2       | 6                        | 6B                     | 2.45  | .36  | --   | --   | --  | .31-0.48  |
| 3       | 5                        | 5B                     | 3.59  | 1.28 | .52  | --   | --  | .42-0.65  |
| 4       | 12                       | 12A                    | 7.59  | 5.21 | 4.05 | .61  | --  | .35-1.54  |
| 1       | 6                        | 6C                     | .87   | --   | --   | --   | --  | .33-1.61  |
| 2       | 6                        | 6B                     | 2.45  | .36  | --   | --   | --  | .31-0.48  |
| 3       | 5                        | 5B                     | 3.59  | 1.28 | .52  | --   | --  | .42-0.65  |
| 4       | 11                       | 11A                    | 7.45  | 5.08 | 3.91 | .52  | --  | .35-0.65  |
| 5       | 1                        | 1A                     | 9.04  | 6.64 | 5.53 | 1.69 | .00 | .00-0.00  |

- Notes: 1. A = below average samples, B = average samples, and C = above average samples.
2. Euclidean distances given are based on input data standardized by subtracting each sample's 10 subtest scores by that variable's grand mean over the 29 samples and dividing by the standard deviation of that variable across samples.

The group subtest profiles of the three sample-groups defined by the 3-cluster solution and of the four sample-groups defined by the 4-cluster solution were subjected to analysis using the MANOVA procedure described in the previous section. There was a significant multivariate main effect for profile elevation,  $F(20,34)=15.9$ ;  $p<.01$ , when the profiles of the groups defined by the 3-cluster solution were tested simultaneously. Examination of the associated 95% confidence interval for each variable for each pair of group differences (see Table 5) reveals that the below average IQ (BA) group profile differs in level from both the average (A) and above average (AA) group profiles on all 10 WISC-R subtests. The A group profile, on the other hand, differs from the AA group profile only on 5 subtests: Information, Similarities, Vocabulary, Comprehension, and Block Design.

TABLE 4

SUBTEST MEANS AND STANDARD DEVIATIONS OF SAMPLE-GROUPS  
DEFINED BY 3- and 4-CLUSTER SOLUTIONS

| Test       | Group BA |      | Group A |      | Group AA |      | Group LA |      | Group HA |      |
|------------|----------|------|---------|------|----------|------|----------|------|----------|------|
|            | Mean     | s.d. | Mean    | s.d. | Mean     | s.d. | Mean     | s.d. | Mean     | s.d. |
| Inf.       | 3.66     | .64  | 9.56    | .95  | 13.85    | 1.55 | 8.64     | .45  | 10.32    | .24  |
| Sim.       | 7.79     | .62  | 9.83    | 1.01 | 14.83    | 1.56 | 9.04     | .90  | 10.48    | .50  |
| Ar.        | 4.23     | .91  | 9.70    | 1.14 | 12.77    | 1.09 | 8.68     | .72  | 10.55    | .53  |
| Voc.       | 3.80     | .80  | 10.15   | 1.03 | 14.37    | 1.57 | 9.34     | .88  | 10.82    | .56  |
| Com.       | 4.81     | .78  | 10.00   | 1.05 | 14.63    | 1.59 | 9.04     | .38  | 10.80    | .63  |
| P.C.       | 5.89     | .94  | 10.31   | .47  | 13.12    | .83  | 9.92     | .15  | 10.63    | .38  |
| P.A.       | 4.38     | .97  | 10.46   | .73  | 13.32    | .86  | 10.06    | .79  | 10.80    | .51  |
| B.D.       | 4.53     | .72  | 10.09   | .81  | 13.40    | 1.17 | 9.38     | .31  | 10.68    | .56  |
| O.A.       | 5.49     | .94  | 10.67   | .66  | 13.47    | 1.21 | 10.40    | .82  | 10.90    | .44  |
| Cod.       | 4.38     | .80  | 9.42    | 1.39 | 11.48    | 1.10 | 8.14     | .98  | 10.48    | .33  |
| Grand Mean | 4.50     | .82  | 10.02   | .58  | 13.52    | 1.02 | 9.26     | .80  | 10.65    | .40  |

There also was a significant multivariate main effect for the group profiles when profile line-segments were tested for parallelism,  $F(18,36) = 6.6$ ;  $p < .01$ . Examination of the associated 95% confidence interval for each profile line-segment for each pair of group differences (see Table 5) reveals that all three groups differ from each other only in the shape of the 4-point profile defined by their Information, Similarities, Arithmetic, and Vocabulary subtest scores. All three groups are similar to each other in having no difference in the slopes of the Picture Completion-Picture Arrangement line-segment and the Object Assembly-Coding line-segment. Group BA differs from group A in the shape of the five Verbal subtest line-segments and is similar in the shape of the five Performance subtest line-segments. Group A is different from group AA in terms of the 4-point profile defined by the Information, Similarities, Arithmetic, and Vocabulary subtests and is similar in respect to the 5-point profile defined by the Performance subtests. The subtest profiles for the three sample-groups defined by the 3-cluster solution are plotted in Figure 1.

For the 4-cluster solution, there was a significant multivariate main effect for profile elevation,  $F(30,48) = 10.1$ ;  $p < .01$ , when the four profiles were tested simultaneously. However, examination of the associated 95% confidence interval for each variable for each pair of group differences (see Table 5) reveals that whereas the below average (BA) group differs on most subtests from the lower average (LA) group there is no difference between the LA group and the higher average (HA) group. Similarly, the HA group differs from the above average (AA) group only on Information, Similarities, Vocabulary, and Comprehension. The multivariate analysis for "elevation" differences between the four groups suggests that groups LA and HA could reasonably be combined to form one larger average group, as suggested by our 3-cluster solution.

TABLE 5

SIGNIFICANT DIFFERENCES BETWEEN GROUPS FOR ELEVATION  
OF SUBTEST SCORES AND FOR SUBTEST-PAIR DIFFERENCES  
(LINE-SEGMENTS) BASED ON SIMULTANEOUS CONFIDENCE  
INTERVALS FROM MANOVA

|                              | Group Comparisons  |       |      |                        |       |       |       |       |       |
|------------------------------|--------------------|-------|------|------------------------|-------|-------|-------|-------|-------|
|                              | 3-Cluster Solution |       |      | 4-Cluster Solution     |       |       |       |       |       |
|                              | BA-A               | BA-AA | A-AA | BA-LA                  | BA-HA | BA-AA | LA-HA | LA-AA | HA-AA |
| Inf.                         | *                  | *     | *    | *                      | *     | *     |       |       | *     |
| Sim.                         | *                  | *     | *    | *                      | *     | *     |       | *     | *     |
| Ar.                          | *                  | *     |      | *                      | *     | *     |       |       |       |
| Voc.                         | *                  | *     | *    | *                      | *     | *     |       |       | *     |
| Com.                         | *                  | *     | *    |                        |       |       |       |       | *     |
| P.C.                         | *                  | *     |      | *                      | *     | *     |       |       |       |
| P.A.                         | *                  | *     |      | *                      | *     | *     |       |       |       |
| B.D.                         | *                  | *     | *    | *                      | *     | *     |       |       |       |
| O.A.                         | *                  | *     |      | *                      | *     | *     |       |       |       |
| Cod.                         | *                  | *     |      | *                      | *     | *     |       |       |       |
| Main effect for elevation:   |                    |       |      |                        |       |       |       |       |       |
| F(20,34) = 15.9; p<.01       |                    |       |      | F(30,48) = 10.1; p<.01 |       |       |       |       |       |
| I-S                          | *                  | *     | *    |                        |       | *     |       |       |       |
| S-A                          | *                  | *     | *    |                        |       | *     |       | *     | *     |
| A-V                          | *                  | *     | *    |                        |       | *     |       | *     |       |
| V-C                          | *                  |       |      | *                      | *     |       |       |       |       |
| C-PC                         |                    | *     | *    | *                      |       | *     |       |       | *     |
| PC-PA                        | *                  | *     |      | *                      | *     | *     |       |       |       |
| PA-BD                        |                    |       |      |                        |       |       |       |       |       |
| BD-OA                        |                    | *     |      |                        |       | *     |       |       |       |
| OA-Co                        |                    |       |      |                        |       |       | *     | *     |       |
| Main effect for parallelism: |                    |       |      |                        |       |       |       |       |       |
| F(18,36) = 6.6; p<.01        |                    |       |      | F(27,50) = 5.2; p<.01  |       |       |       |       |       |

Note: an \* denotes significance with alpha = .05.

There was also a significant multivariate main effect for the four group-profiles when profile line-segments were tested for parallelism,  $F(27,50)=5.2; p<.01$ . Again, however, examination of the associated 95% confidence interval for each profile line-segment for each pair of group differences (see Table 5) reveals little difference between the shapes of the profiles for groups LA and HA. These two groups differ solely in respect to the slope of the line-segment between Object Assembly and Coding. The subtest profiles for the two average groups defined by the 4-cluster solution are plotted in Figure 2.

Generally, the MANOVA procedures carried out support a 3-group interpretation of the 29-sample data with the resulting three profiles demonstrating significant differences in both score elevation and profile shape or subtest patterning.



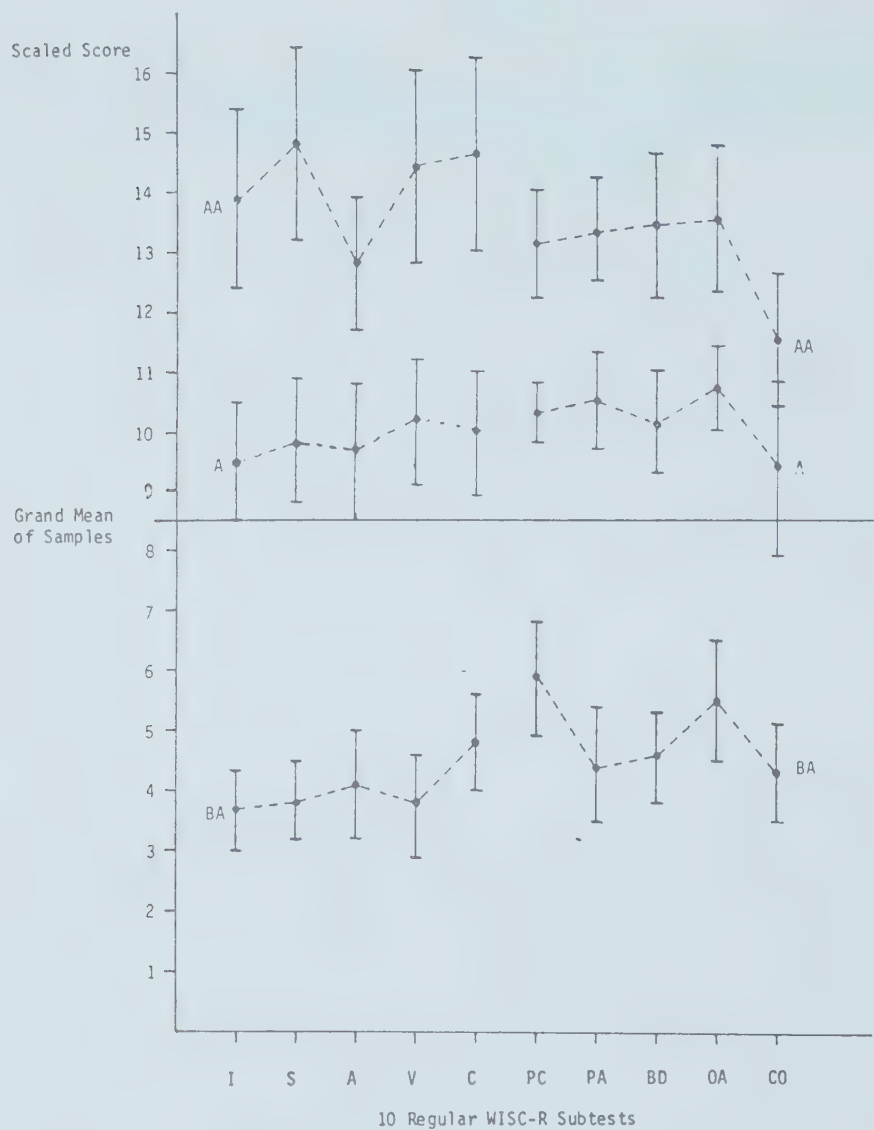


Figure 1. Mean subtest scores and S.D.s for below average (BA), average (A), and above average (AA) I.Q. clusters.

Discussion

Cluster analysis and MANOVA procedures applied to the WISC-R subtest profiles of 29 samples of children representing three levels of overall intellectual functioning revealed three relatively distinct performance patterns; one for below average IQ children, one for average IQ children, and one for above average IQ children. The subtest profiles that emerged for the below average and above average IQ groups of samples of this study are in reasonable agreement with previous research findings for retarded and bright populations (Kaufman & Van Hagen, 1977; Lutey, 1977; Naglieri, 1980; Schiff et al., 1981).

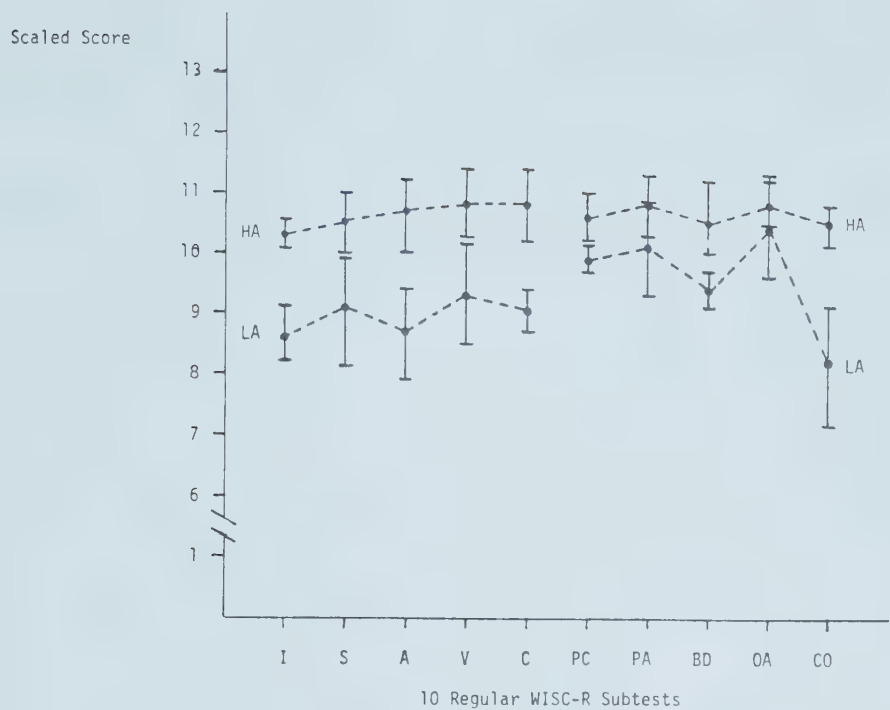


Figure 2. Mean subtest scales scores and S.D.s for lower average (LA) and higher average (HA) I.Q. clusters.

As a group, the below average IQ samples of this study tended to score lowest relative to their own mean performance level on three of the four WISC-R subtests which together form Kaufman’s (1975) *Verbal Comprehension* factor. Also, as might be expected, this group demonstrated relatively poor performance on the four subtests—Information, Similarities, Vocabulary, and Block Design—which have been reported to load most heavily and consistently on a general intelligence factor (Lutey, 1977; Wallbrown et al., 1975). On the other hand, the below average samples showed relatively good performance on two of the four subtests—Picture Completion and Object Assembly—that load on Kaufman’s *Perceptual Organization* factor. It appears that below average IQ children might be best characterized by a subtest profile of highest scores on the Picture Completion and Object Assembly subtests, middle scores on Vocabulary and Block Design, and lowest scores on Similarities and Information. Of the 12 samples of below average children included in our meta-analysis, 100% had Picture Completion rank 1 or 2 (highest score=1; lowest=10) and 83% had Object Assembly rank 1 or 2. Similarly, in 75% of the below average samples Similarities or Information ranked 9 or 10. Coding or Picture Arrangement ranked in the middle (5 or 6) in 67% of these samples.

In contrast to the below average group, the present analysis found the above average samples to score relatively high on all four of the WISC-R subtests which make up the *Verbal Comprehension* factor and to do less well on the four subtests which make up the *Perceptual Organization* factor. Above average IQ children appear to be best characterized by a subtest profile of highest scores on the Comprehension and Similarities subtests, middle scores on Object Assembly and Picture Arrangement, and lowest scores on Arithmetic and Coding. Of the 6 above

average samples included in this study, 100% had Comprehension or Similarities rank 1 or 2, 83% had Object Assembly or Picture Arrangement rank 5 or 6, 83% had Arithmetic rank 9, and 100% had Coding rank 10. Compared to the below average IQ group of samples, not one of the above average samples had Picture Completion rank first or Similarities or Information rank last.

Unique to the present study was an examination of WISC-R subtest patterning across a number of normal average IQ samples and, as might be expected given the results of Kaufman's (1976a, 1976b, 1976c) research on WISC-R subtest scatter in the standardization sample, the present analysis revealed average children to have a reasonably consistent and stable ordering of WISC-R subtest scores rather than a "flat" profile. Although it must be noted that the average samples showed less consistency in their profiles than did either the below average or above average samples, the average group of samples generally demonstrated a subtest score pattern that reflected aspects of the patterns of both the below average group and the above average group. Like the below average group, the average group tended to perform best on subtests that load on the *Perceptual Organization* factor of the WISC-R whereas, like the above average, they performed more poorly on the subtests which load the *Freedom from Distractibility* factor. Of the 11 samples which formed our average IQ group, 64% had Object Assembly or Picture Arrangement rank 1 or 2, 91% had Vocabulary or Block Design rank 5 or 6, and 91% had Arithmetic, Information, or Coding rank 9 or 10. Compared to the below average group where Picture Completion ranked 1 or 2 in 100% of the samples and Coding ranked 9 or 10 in 0%, for the average group Picture Completion ranked 1 or 2 in only 18% and Coding ranked 9 or 10 in 45%.

Interestingly, although the data from our samples did not include the Digit Span subtest, our average IQ sample tended to perform most poorly on three (Arithmetic, Coding, Information) of the four subtests which together comprise the *ACID* group of subtests (Arithmetic, Coding, Information, Digit Span) which have been implicated in the diagnosis of learning disabilities (see Kaufman, 1979a, 1980, 1981). Just as do the learning disabled, our samples of normal average IQ children also appear to find the *ACID* subtests relatively difficult. Possibly the "learning disabilities profile" suggested by a number of researchers as being characteristic of LD children (e.g., Clarizio & Bernard, 1981; Rugel, 1974; Smith, et al., 1977; Vance & Singer, 1979) is a more general characteristic of average IQ children.

In summary, although the WISC-R factor structure has been shown to be relatively invariant across IQ levels, the relative patterning of WISC-R subtests in children's profiles would appear to be related in some part to the overall level of the child's intellectual functioning. Such subtests as Information, Similarities, and Vocabulary tend to come earlier in ranking with increasing FSIQ while such subtests as Picture Completion, Arithmetic, and Coding tend to come later with increasing FSIQ. Retarded children demonstrate a distinctly different pattern of subtest performance than do average IQ children who, in turn, demonstrate a pattern that is different from that of the intellectually superior. In terms of the three-factor breakdown of WISC-R subtests suggested by Kaufman (1975), it appears that low IQ children may be characterized by the pattern *Perceptual Organization* > *Freedom from Distractibility* > *Verbal Comprehension*, that average IQ children may be characterized by the pattern *Perceptual Organization* > *Verbal Comprehension* > *Freedom from Distractibility*, and that bright children may be characterized by the pattern *Verbal Comprehension* > *Perceptual Organization* > *Freedom from Distractibility*.



Although the findings of this study should not come as a great surprise to any researcher or clinician familiar with the WISC-R and its legitimate uses, these results have important implications for future research aimed at determining characteristic WISC-R performance profiles for various diagnostic or exceptional groups. In future, such studies will need to demonstrate that the WISC-R performance pattern evidenced by a particular diagnostic group differs not only from a flat profile with a mean scaled score of 10 points, but also from the profile common to normal children of the same intellectual level. Certainly, a comparison of the characteristic profile demonstrated by many samples of learning disabled children to that of the average IQ group of this study would greatly decrease researchers' confidence in the utility of the WISC-R for differentiating learning disabled from normal children (see Berk, 1983; Miller & Walker, 1981).

While the three group profiles that have emerged from the present analysis are probably indicative of a specific trend of relative performance within the individual children comprising each group of samples, such an overall trend may or may not be consistent with the profile shown by any one sample or any individual child within a sample (Berk, 1983; Huelsman, 1970). At best, the present results are likely to be only generalizable to large groups of children of fairly homogeneous intellect. The authors do not suggest WISC-R group profiles be applied to the differential diagnosis of individual children; rather such group-profiles are more likely to be useful at a more molar level in suggesting broad cognitive differences between levels of intellectual functioning. In a general sense, WISC-R profiles are more likely diagnostic of intellectual-level differences than more subtle psycho-emotional differences within any one level of general intellect.

The possibility also exists that the profile differences among groups may be a function of the intercorrelations among subtests. Since the groups formed are differentiated on FSIQ, which is a composite of the 10 WISC-R subtests, subtests that are highly correlated with FSIQ would be expected to differentiate these groups in a similar manner. Subtests with low correlations with FSIQ would be expected to be less successful in differentiating such groups. Thus groups scoring at the two extremes of FSIQ would be expected to have mirror-image profiles. The authors are planning further study into profile differences as a function of the various subtests' squared multiple correlation with the remaining tests.

Finally, it is suggested that a meta-analysis of the WISC-R subtest data from all available research studies of retarded, learning disabled, normal, emotionally disturbed, delinquent, ethnic minorities, and gifted children would be invaluable in helping to settle the question of whether or not characteristic WISC-R profiles exist for groups defined by non-intellective criteria. It is the authors' contention that such a study will likely find that all diagnostic groups, except for the seriously brain-injured, will show patterns of WISC-R subtest performance similar to that of normal children of comparable general intelligence.

#### Notes

1. Requests for reprints should be addressed to Horst H. Mueller at the Department of Educational Psychology, Faculty of Education, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.
2. This study was partially supported by the Social Sciences and Humanities Research Council of Canada through a Doctoral Fellowship to H. Mueller, and a University of Alberta Central Research Fund grant to R. Short.
3. The authors wish to acknowledge and thank Dr. Steven S. Dennis, Department of Occupational Therapy, University of Alberta, for his invaluable assistance in the preparation of this manuscript.

References

- Alper, A. (1967). An analysis of the WISC with institutionalized retardates. *American Journal of Mental Deficiency*, 71(6), 624-630.
- Anderson, M., Kaufman, A., & Kaufman, N. (1976). Use of the WISC-R with a learning disabled population: Some diagnostic implications. *Psychology in the Schools*, 13(3), 381-386.
- Bannatyne, A. (1974). Diagnosis: A note on recategorization of the WISC-R scaled scores. *Journal of Learning Disabilities*, 7(2), 272-274.
- Baumeister, A. (1964). The use of the WISC with mental retardates: A review. *American Journal of Mental Deficiency*, 69(2), 183-194.
- Berk, R. (1983). The value of WISC-R profile analysis for differential diagnosis of learning disabled children. *Journal of Clinical Psychology*, 39(1), 133-136.
- Catron, D., & Catron, S. (1977). WISC-R vs. WISC: A comparison with educable mentally retarded children. *Journal of School Psychology*, 15(3), 264-266.
- Clarizio, H., & Bernard, R. (1981). Recategorized WISC-R scores of LD children and differential diagnosis. *Psychology in the Schools*, 18(1), 5-12.
- Covin, T. (1977). Stability of the WISC-R for 9-year-olds with learning difficulties. *Psychological Reports*, 40, 382.
- Cronbach, L., & Gleser, G. (1953). Assessing similarity between profiles. *Psychological Bulletin*, 6(3), 456-473.
- Dean, R. (1977). Patterns of emotional disturbance on the WISC-R. *Journal of Clinical Psychology*, 33(3), 486-490.
- Dean, R. (1978). Distinguishing learning disabled and emotionally disturbed children on the WISC-R. *Journal of Consulting & Clinical Psychology*, 46(2), 381-382.
- Dixon, W. (Ed.) (1981). *BMDP Statistical Software, 1981 ed.* Los Angeles, CA: University of California Press.
- Dudley-Marling, C., & Kaufman, N. (1981) WISC & WISC-R profiles of LD children: A review. *Learning Disabilities Quarterly*, 4(3), 307-319.
- Finn, J. (1974). *A general model for multivariate analysis*. New York, N.Y.: Holt, Rinehart & Winston.
- Gironda, R. (1977). A comparison of WISC and WISC-R results of urban EMR students. *Psychology in the Schools*, 14(2), 271-275.
- Glass, G. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3-8.
- Gutkin, T. (1979). WISC-R scatter indices: Useful information for differential diagnoses? *Journal of School Psychology*, 17(4), 368-371.
- Gutkin, T., & Reynolds, C. (1980). Factorial similarity of the WISC-R for Anglos and Chicanos referred for psychological services. *Journal of School Psychology*, 18(1), 34-39.
- Hale, R., & Landino, S. (1981). Utility of WISC-R subtest analysis in discriminating among groups of conduct problem, withdrawn, mixed, and nonproblem boys. *Journal of Consulting & Clinical Psychology*, 49(1), 91-95.
- Hale, R., & Saxe, J. (1983). Profile analysis of the WISC-R. *Journal of Psychoeducational Assessment*, 1(2), 155-162.
- Hamm, H., & Evans, J. (1978). WISC-R subtest patterns of severely emotionally disturbed students. *Psychology in the Schools*, 15(1), 188-190.
- Hays, J., & Solway, K. (1977). Violent behavior and differential WISC-R characteristics. *Journal of Consulting & Clinical Psychology*, 45(6), 1187.
- Hays, J., Solway, K., & Schreiner, D. (1978). Intellectual characteristics of juvenile murderers versus status offenders. *Psychological Reports*, 43(1), 80-82.

- Henry, S., & Wittman, R. (1981). Diagnostic implications of Bannatyne's recategorized WISC-R scores for identifying learning disabled children. *Journal of Learning Disabilities*, 14(9), 517.
- Hirshoren, A., & Kavale, K. (1976). Profile analysis of the WISC-R: A continuing malpractice. *Exceptional Children*, 23(1), 83-87.
- Huelsman, C. (1970). The WISC-subtest syndrome for disabled readers. *Perceptual and Motor Skills*, 30, 535-550.
- Hunka, S. (1980). One-way multivariate analysis of variance and covariance (MULV16). Unpublished program documentation. Division of Educational Research Services, Faculty of Education, University of Alberta, Edmonton, Alberta.
- Karnes, F., & Brown, K. (1980). Factor analysis of the WISC-R for the gifted. *Journal of Educational Psychology*, 72(2), 197-199.
- Kaufman, A. (1975). Factor analysis of the WISC-R. *Journal of Consulting & Clinical Psychology*, 43(1), 135-147.
- Kaufman, A. (1976a). A new approach to the interpretation of test scatter on the WISC-R. *Journal of Learning Disabilities*, 9(3), 160-167.
- Kaufman, A. (1976b). Verbal-Performance IQ discrepancies on the WISC-R. *Journal of Consulting & Clinical Psychology*, 44(5), 739-744.
- Kaufman, A. (1976c). Do normal children have "flat" ability profiles? *Psychology in the Schools*, 13(2), 284-285.
- Kaufman, A. (1979a). *Intelligent Testing With the WISC-R*. New York, N.Y.: J. Wiley & Sons.
- Kaufman, A. (1979b). WISC-R research: Implications for interpretation. *School Psychology Digest*, 8(1), 5-27.
- Kaufman, A. (1979c). The role of speed on WISC-R performance across the age range. *Journal of Consulting & Clinical Psychology*, 47(3), 595-597.
- Kaufman, A. (1980). Issues in psychological assessment: Interpreting the WISC-R intelligently. In B. Lahey & A. Kazdin (Eds.), *Advances in clinical child psychology*. New York, N.Y.: Plenum.
- Kaufman, A. (1981). The WISC-R and learning disabilities assessment: State of the art. *Journal of Learning Disabilities*, 14(9), 520-526.
- Kaufman, A., & Van Hagen, J. (1977). Investigation of the WISC-R for use with retarded children: Correlations with the 1972 Stanford-Binet and comparison of WISC and WISC-R profiles. *Psychology in the Schools*, 14(1), 10-14.
- Kunce, J., & McMahon, R. (1979). Neuropsychological significance of children's Wechsler intelligence scores. *Psychological Reports*, 44(3, pt.1), 787-790.
- Lutey, C. (1977). *Individual intelligence testing: A sourcebook*. Greeley, Colorado: Lutey Publishing.
- Miller, M. (1980). On the attempt to find WISC-R profiles for learning and reading disabilities. *Journal of Learning Disabilities*, 13(6), 52-54.
- Miller, M., & Walker, K. (1981). The myth of the LD WISC-R profile. *Exceptional Child*, 28(2), 83-88.
- Moore, D., & Wielen, O. (1981). WISC-R scatter indexes of children referred for reading diagnosis. *Journal of Learning Disabilities*, 14(9), 511-516.
- Morris, J., Evans, J., & Pearson, D. (1978). The WISC-R subtest profile of a sample of severely emotionally disturbed children. *Psychological Reports*, 42(3), 319-325.
- Morrison, D. (1967). *Multivariate statistical methods*. New York, N.Y.: McGraw-Hill.
- Mueller, H. (1979). WISC-R subtest patterns of retarded and bright children. Unpublished manuscript available from the author, Department of Educational Psychology, University of Alberta, Edmonton, Alberta.
- Mueller, H., Mancini, G., & Short, R. (in press). An evaluation of the diagnostic efficiency of the WISC-R. *Alberta Journal of Educational Research*.



- Nagle, R., & Lazarus, S. (1979). The comparability of the WISC-R and WAIS among 16-year-old EMR children. *Journal of School Psychology, 17*(4), 362-367.
- Naglieri, J. (1980). WISC-R subtest patterns for LD and retarded children. *Perceptual & Motor Skills, 51*(2), 605-606.
- Ollendick, T. (1979). Discrepancies between Verbal and Performance IQs and subtest scatter on the WISC-R for juvenile delinquents. *Psychological Reports, 45*(4), 563-568.
- Peters, H. (1976). The validity of the WISC-R. *Canadian Journal of Behavioral Science, 8*(4), 414-417.
- Pristo, L. (1978). Comparing WISC and WISC-R scores. *Psychological Reports, 42*(4), 515-518.
- Quattrocchi, M., & Sherrets, S. (1980). WISC-R: The first five years. *Psychology in the Schools, 17*(3), 297-312.
- Raviv, A., Margalith, M., Raviv, A., & Sade, E. (1981). The cognitive patterns of Israeli LD children as reflected in the Hebrew version of the WISC-R. *Journal of Learning Disabilities, 14*(7), 411-415.
- Reschley, D., & Davis, R. (1977). Comparability of WISC and WISC-R scores among borderline and mildly retarded children. *Journal of Clinical Psychology, 33*(4), 1045-1047.
- Resnick, R. (1977). An abbreviated form of the WISC-R: Is it valid? *Psychology in the Schools, 14*(4), 426.
- Rowe, H. (1977). Borderline versus mentally deficient: A study of the performance of educable mentally retarded adolescents on WISC-R and WISC. *Australian Journal of Mental Deficiency, 4*(5), 11-14.
- Rugel, R. (1974). WISC subtest scores of disabled readers. *Journal of Learning Disabilities, 7*(1), 57.
- Ryckman, D. (1981). Searching for a WISC-R profile for learning disabled children: An inappropriate task? *Journal of Learning Disabilities, 14*(9), 507-526.
- Sandoval, J. (1982). The WISC-R factorial validity for minority groups and Spearman's hypothesis. *Journal of School Psychology, 20*(3), 198.
- Sattler, J. (1982). *Assessment of children's intelligence and special abilities, 2nd Edition*. Boston, Mass.: Allyn & Bacon.
- Schiff, M., Kaufman, A., & Kaufman, N. (1981). Scatter analysis of WISC-R profiles for LD children with superior intelligence. *Journal of Learning Disabilities, 14*(7), 400-404.
- Schooler, D., Beebe, M., & Koepke, T. (1978). Factor analysis of the WISC-R scores for children identified as learning disabled, educable mentally retarded, and emotionally impaired. *Psychology in the Schools, 15*(4), 478-485.
- Schwarting, F. (1976). A comparison of the WISC and WISC-R. *Psychology in the Schools, 13*(1), 139.
- Silverstein, A. (1968). WISC subtest patterns of retardates. *Psychological Reports, 23*(4), 1061-1062.
- Simensen, R., & Sutherland, J. (1974). Psychological assessment of brain damage: The Wechsler scales. *Academic Therapy, 10*(1), 69-81.
- Smith, M., Coleman, J., Dokecki, P., & Davis, E. (1977a). Intellectual and academic characteristics of school-verified LD children. *Exceptional Child, 43*, 352-357.
- Smith, M., Coleman, J., Dokecki, P., & Davis, E. (1977b). Recategorized WISC-R scores of LD children. *Journal of Learning Disabilities, 10*(7), 437-443.
- Stokes, E., Brent, D., Huddleston, N., & Rozier, J. (1978). Comparison of WISC and WISC-R scores of sixth grade students: Implications for validity. *Educational & Psychological Measurement, 38*(3), 469-473.
- Swerdlik, M., & Wilson, F. (1979). A comparison of WISC and WISC-R subtest scatter. *Journal of Learning Disabilities, 12*(1), 46-48.
- Tew, B. (1977). Spina Bifida children's scores on the WISC. *Perceptual & Motor Skills, 44*(2), 381-382.

- Thompson, J., & Finley, C. (1962). A further comparison of the intellectual patterns of gifted and mentally retarded children. *Exceptional Children*, 28(2), 379.
- Van Hagen, J., & Kaufman, A. (1975). Factor analysis of the WISC-R for a group of mentally retarded children and adolescents. *Journal of Consulting & Clinical Psychology*, 43(4), 661-667.
- Vance, H. (1979). Sex differences on the WISC-R for retarded children and youth. *Psychology in the Schools*, 16(1), 27-31.
- Vance, H., Gaynor, P., & Coleman, M. (1976). Analysis of cognitive abilities for learning disabled children. *Psychology in the Schools*, 13(4), 477-483.
- Vance, H., Gaynor, P., & Coleman, M. (1977). Item analysis of the WISC-R. *Psychology in the Schools*, 14(1), 132-139.
- Vance, H., Hankins, N., Wallbrown, F., Engin, A., & McGee, H. (1978). An analysis of cognitive abilities for mentally retarded children on the WISC-R. *Psychological Record*, 28(3), 391-397.
- Vance, H., & Singer, M. (1979). Recategorization of the WISC-R subtest scaled scores for learning disabled children. *Journal of Learning Disabilities*, 12(8), 487-490.
- Wallbrown, F., Blaha, J., Wallbrown, J., & Engin, A. (1975). The hierarchical factor structure of the WISC-R. *Journal of Psychology*, 89(2), 223-235.
- Webster, R., & Lafayette, A. (1980). Distinguishing among three subgroups of handicapped students using Bannatyne's recategorization of the WISC-R. *Journal of Educational Research*, 73(4), 237-240.
- Wechsler, D. (1974). *WISC-R Manual*. New York, N.Y.: Psychological Corporation.
- Wersh, J., & Briere, J. (1981). WISC-R subtest variability in normal Canadian children and its relationship to sex, age, and IQ. *Canadian Journal of Behavioral Science*, 13(1), 77-86.
- Wheaton, P., & Vandergriff, A. (1978). Comparison of WISC and WISC-R scores of highly gifted students in public school. *Psychological Reports*, 43(4), 627-630.

## List of Publications Received

- Anderson, L. & Windham, D. M. (Eds.) *Education and development*. Toronto: Lexington Books, D. C. Heath Canada Ltd., 1982, 213 pp.
- Axelrod, P. *Scholars and dollars. Politics, economics, and the universities of Ontario 1945-1980*. Toronto: University of Toronto Press, 1982, 271 pp., \$13.95 (paper), \$35.00 (cloth).
- Baine, D. *Instructional design for special education*. Englewood Cliffs, NJ: Educational Technology Publications, 1982, 321 pp., \$24.95.
- Barrett, M. *Population and Canada. The social problems in Canada Series 6*. Toronto: Faculty of Education Guidance Centre, University of Toronto. University of Toronto, 1982.
- Barrett, H. O. *English. (The student, subject, and careers series)*. Toronto: Guidance Centre,
- Batten, M., & Girling-Butcher, S. *Perceptions of the quality of school life*. Victoria, Australia: Australian Council for Educational Research Ltd., ACER Monograph No. 12, 1981, 64 pp., \$6.50.
- Biggs, J. B., & Collis, K. F. *Evaluating the quality of learning*. New York: Academic Press Inc., 1981, 245 pp., \$26.50.
- Bowd, A., McDougall, D., & Yewchuk, C. Ed. *Psych. A Canadian perspective*. Toronto: Gage Publishing Limited, 1982, 470 pp., \$21.95.
- Brigham Young University (Ed.) *How to involve parents in early childhood education*. Provo, Utah: 1982, 154 pp.
- Canadian Education Association, Toronto, Ontario:
- Requirements for teaching certificates in Canada*, September, 1982, \$5.00.
- Strategies for public involvement: Final report of the CEA Task Force on Public Involvement in Educational Decisions*, September, 1981, 104 pp., \$6.00.
- Chilver, P., & Gould, G. *Learning and language in the classroom*. New York: Pergamon Press, 1982, \$19.90, 151 pp.
- Crespo, M., & Haché. *Gestion et décroissance en éducation. Le cas d'une commission scolaire québécoise*. Montreal: Les Presses de l'Université de Montréal, 1983, 139 pp., \$15.00.
- Czumer, W. A. *Recollections about the life of the first Ukrainian settlers in Canada*. Downsview, Ontario: University of Toronto Press, 1981, 176 pp., \$5.95 (paper).
- Fraser, I., & Brownhill, K. *Reading is not for me*. Toronto: Clarke, Irwin & Company Ltd., 1983, 208 pp., \$9.95.
- French, E. *The promotion of literacy in South Africa*. Pretoria: Human Sciences Research Council, Institute for Research into Language and the Arts, 1982.
- Friesen, J. W. *Schools with a purpose*. Calgary: Detselig Enterprises Ltd., 1983, 142 pp.
- Garbarino, J., & Asp, C. E. *Successful schools and competent students*. Toronto: Lexington Books, D. C. Heath, 1981, 170 pp., \$24.95.
- Gouvernement du Québec, Ministère de l'éducation:
- Elèves, étudiants, étudiantes: Présence active et responsable. Orientation et voies d'action*. 1982, 31 pp.
- Etude des arrivées chez le personnel enseignant du réseau public du Québec, de 1971-72 à 1977-78*. Louis Dionne, Mars 1981.
- Inventaire des travaux publiés par le service de la démographie scolaire au 1 avril 1981*. Desmond Dufour, avril 1981.



*Les clientèles universitaires de 1973 à 2001.* Document 59. André Lespérance, 1981.

*Mother tongue and language of instruction in Québec's public schools: Recent evolution.*

*Perspectives de clientèles scolaires pour le Québec et ses régions administratives scolaires 1981 à 2001, selon l'âge, le sexe, et le niveau d'enseignement.* Desmond Dufour, février 1981.

Government of Canada, Ministry of State for Science and Technology:

*Federal science activities, 1982/83.*

MOSST Background Paper No. 14. *Recent trends in degrees awarded and enrolments at Canadian universities.* 68 pp.

MOSST Background Paper No. 15. *University enrolment projections to 2000.* 60 pp.

MOSST Background Paper No. 16. *The stock of research trained personnel.* 22 pp.

MOSST Background Paper No. 17. *Research manpower requirements arising from accelerated expenditures on R & D.* 30 pp.

MOSST Background Paper No. 18. *The requirements for engineering graduates to 1985,* 1981.

MOSST Background Paper No. 20. *The Canadian space program plan for 1982/83 - 1984/85,* December 1981.

Government of Canada, Statistics Canada, Education, Science and Culture Division:

*Elementary-secondary school enrolment, 1980-81.* Catalogue No. 81-210.

*Financial statistics of education, 1977-78.* Catalogue No. 81-208.

*Financial statistics of education, 1978-79.* Catalogue No. 81-208.

*Salaries and qualifications of teachers in public elementary and secondary schools, 1980-81.* Catalogue No. 81-202.

*Teachers in universities, 1979-80.* Catalogue No. 81-241.

*Teachers in universities, 1980-81.* Catalogue No. 81-241.

Gribbons, W. D., & Lohnes, P. R. *Careers in theory and experience. A twenty-year longitudinal study.* Albany: State University of New York Press, 1982, 145 pp., \$14.95 (paper).

Gross, F. L., Jr. *Passages in teaching. Predictable crises in the teaching of adolescents and young adults.* New York: Philosophical Library Publishers, 1982, 221 pp., \$15.00.

Hale, J. E. *Black children: Their roots, culture, and learning styles.* Provo, Utah: Brigham Young University Press, 1982, 191 pp., \$9.95.

*Investigaciones terminadas o iniciadas durante 1981-82.* Universidad de Los Andes.

Jungo, M. E. (Compiled by). *International bibliography for a didactics of early bilingualism in the education of underprivileged children, especially children of migrant workers.* Einsiedeln, Switzerland, 1982.

- Lombard, A. D. *Success begins at home*. Toronto: Lexington Books, D. C. Heath & Co., 150 pp., \$22.50.
- Maier, N. (Ed.). *Teaching the gifted, challenging the average*. Toronto: Faculty of Education Guidance Centre, 1982, 122 pp.
- Manley-Casimir, M. E. *Family choice in schooling*. Toronto: Lexington Books, D. C. Heath & Co., 210 pp., \$28.75.
- McGowan, C. *In the beginning . . . A scientist shows why the creationists are wrong*. Toronto: Macmillan, 1983, 256 pp., \$18.95 (cloth).
- McPeck, J. *Critical thinking and education*.
- Miffen, F. J., & Miffen, S. C. *Sociology of education: Canada and beyond*. Calgary: Detselig Enterprises Ltd., 1982, 407 pp., \$18.95 (paper).
- Neufeldt, R. W. *Religious studies in Alberta: A state-of-the-art review*. Calgary: University of Calgary, 1983, 145 pp.
- New Zealand Council for Educational Research. *Forty-seventh annual report, 1981-82*. NZCER, 178 Willis St., Wellington.
- Ontario Ministry of Education. *Learning abilities: Identification and intervention practices*. Principal investigators: I. Davidson, H. Silverman, M. Hughes. Toronto, 1981, 122 pp., \$5.00.
- Patton, P. C., & Holoien, R. A. *Computing in the humanities*. Toronto: Lexington Books, D. C. Heath Canada Ltd., 1981, 404 pp., \$37.50.
- Phipp, J. *The electronic classroom*. A CEA report. Toronto: Canadian Education Association, 39 pp., 1982.
- Quiring, I. *Strangled roots*. Calgary: Detselig Enterprises Ltd., 1982, 186 pp.
- Roberts, H. *Culture and adult education*. A study of Alberta and Quebec. Edmonton: University of Alberta Press, 1982, 274 pp.
- Roe, M. *Multiculturalism, racism, and the classroom*. A CEA report. Toronto: Canadian Education Association, 1982, 68 pp., \$6.00.
- Rooke, P. T., & Schnell, R. L. *Studies in childhood history. A Canadian perspective*. Calgary: Detselig Enterprises Ltd., 1982, 212 pp.
- Ross, J. A., & Maynes, F. *Geographic thinking skills*. Toronto: Ontario Institute for Studies in Education, 1981, 198 pp., \$10.00.
- Ross, J. A., & Maynes, F. J. *Teaching problem-solving*. Toronto: Ontario Institute for Studies in Education Press, 1982, 208 pp.
- Rowntree, D. *Educational technology in curriculum development* (Second Ed.). Troy, Missouri: Ballinger Publishing Co., 196 pp.
- Scholer, M. *La technologie de l'éducation. Concept, bases et application*. Montreal: Presses Universitaires de l'Université de Montréal, 1983, 197 pp., \$25.50.
- Scottish Council for Research in Education. *Fifty-third annual report, 1980-81*. 72 pp.
- Solmon, L. C., Kent, L., Oshner, N. L., & Hurwicz, M. *Under-employed Ph.D.'s*. Toronto: Lexington Books, D. C. Heath & Co., 1981, 350 pp., \$40.95.
- Stamp, R. M. *The schools of Ontario 1876-1976*. Toronto: University of Toronto Press, 1982, 293 pp., \$30.00.
- Stanworth, M. *Gender and schooling A study of sexual divisions in the classroom*. Toronto: Copp Clark Pitman, 1983, 64 pp.
- Stewart, F. K. *The Canadian Education Association 1957-1977*. Toronto: CEA, 1982, 131 pp., \$8.00.
- Tarnopol, L., & Tarnopol, M. *Comparative reading and learning difficulties*. Toronto: Lexington Books, D. C. Heath & Co., 1981, 562 pp., \$45.95.
- Tattum, D. *Disruptive pupils in schools and units*. New York: John Wiley & Sons, 1982, 329 pp., \$32.95 (cloth):

- Titley, E. B., & Miller, P. J. *Education in Canada: An interpretation*. Calgary: Detselig Enterprises Ltd., 1982, 228 pp., \$16.95 (paper).
- Travers, R. M. W. *How research has changed American schools. A history from 1840 to the present*. Kalamazoo, MI: Mythos Press, 1983, 611 pp., \$25.00 (cloth).
- Waterhouse, L. H., Fischer, K. M., & Bouchard Ryan, E. *Language awareness and reading*. Newark, DE: International Reading Association, 1980, 59 pp.
- Watson, J. *Breaking the blame chain. An informal observation report on the teaching and learning of English as a first language*. Dubuque, Iowa: Kendall/Hunt Publishing Co., 1981, 29 pp.
- Weaver, W. T. *The contest for educational resources*. Toronto: University of Toronto Press, 1982, 188 pp., \$28.75.
- Williams, T., & Batten, M. *The quality of school life*. Victoria, Australia: Australian Council for Educational Research Ltd., ACER Monograph No. 12, 1981, 64 pp., \$6.50.
- Wilson, J. D. *Canadian education in the 1980s*. Calgary: Detselig Enterprises Ltd., 1981, 282 pp., \$13.95.
- Willie, C. V. *The ivory & ebony towers*. Toronto: D. C. Heath Canada Ltd., 1981, 173 pp., \$24.95.



# THE JOURNAL OF EDUCATIONAL THOUGHT

Published three times yearly by the Faculty of Education  
The University of Calgary

Telephone (403) 284-5629

The Journal of Educational Thought promotes speculative, critical and historical research concerning the theory and practice of education in a variety of areas: administration, comparative education, curriculum, educational communications, evaluation, instructional methodology, intercultural education, philosophy, psychology, and sociology. Critical and controversial views on the field of education or on previously published articles in The Journal of Educational Thought appear in a section entitled The Forum, which also contains papers of a less formal nature. The book section includes short book reviews as well as extended book review essays. The Journal is international in scope, qualitative in nature, and serves a broad readership; specialists in the areas mentioned, scholars, and the public in general.

---

**Instructions for the preparation of manuscripts appear on the last page of each issue.**

---

**Publishers are invited to submit to the Editors books for review.**

---

The Journal is indexed in *Canadian Education Index* and by the *Educational Resources Information Centre (ERIC)*. Abstracts of appropriate articles appear in *Educational Administration Abstracts*, *Language and Language Behaviour Abstracts*, and *Sociology of Education Abstracts*. Available on microfilm from University Microfilms, Inc., Ann Arbor, Michigan.

It is the policy of the Journal to provide an open forum for scholarly discussion. Editorials as well as articles are to be regarded as expressing the viewpoints of the writers themselves, and not as reflecting official positions of the Editorial Committee, the Faculty of Education, or The University of Calgary. It is an additional policy of *THE JOURNAL OF EDUCATIONAL THOUGHT* to modify at its discretion wordy and sexist expressions appearing in a manuscript submitted for publication, whenever a change in these expressions will not distort the meaning of the passage.

---

## Subscription Rate

Effective January 1st, 1982

\$15.00 - 1 year (individual)  
\$20.00 - 1 year (institutions)  
\$ 6.00 - single issue  
\$ 2.00 - back issue  
\$ 7.00 - Special issue

## Address

The Journal of Educational Thought  
Faculty of Education  
The University of Calgary  
2500 University Drive N.W.  
Calgary, Alberta, Canada  
T2N 1N4

Postage overseas air printed (not U.S.), additional \$2.00.

**Claims for undelivered copies must be made within one month following the month of publication.**











## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by authors' last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary* or to the *Dictionary of Canadian English*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.





# ajer

THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH

VOLUME XXX

NUMBER 2

JUNE 1984

PUBLISHED BY  
THE UNIVERSITY OF ALBERTA • EDMONTON

# THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

Published in March, June, September, December by the  
Faculty of Education, The University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

EDITOR: C. S. Bumbarger

## CONSULTING EDITORS

D. C. Bjorkquist  
University of Minnesota

P. E. Bryant  
University of Oxford

B. M. Chung  
Chungbuk University,  
Republic of Korea

A. Côté  
University of Ottawa

D. Fitzgerald  
The University of New England

D. B. Harrison  
The University of Calgary

N. L. Hersom  
The University of Saskatchewan

M. Laferrière  
McGill University

T. Morrison  
University of Manitoba

J. Richardson  
University of New Brunswick

P. Robinson  
Atlantic Institute of Education

T. Rogers  
University of British Columbia

R. E. Traub  
Ontario Institute for  
Studies in Education

K. Watson  
University of Reading

R. J. Wilson  
Queen's University

*The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, *Social Science Citation Index*, *ERIC/Current Index to Journals in Education*, and *Alberta Education Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, *Sociology of Education Abstracts*, *Language Behavior Abstracts*, *Cognitive Development Abstracts*, *Research into Higher Education Abstracts*, *School Organization and Management Abstracts*, and *Multicultural Education Abstracts*.

The subscription rate is \$12.00 per year; single copies are \$3.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, The University of Alberta, Edmonton, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436





# The Alberta Journal of Educational Research

Volume XXX, Number 2

June, 1984

CONTENTS

|  |     |
|--|-----|
| The Effect of Homogeneous vs. Heterogeneous Matching-Item Format<br>on Test Performance and Reliability .....                | 91  |
| <i>D. E. Allison</i>   |     |
| A Detailed Analysis of Group Differences on the California<br>Short-Form Test of Mental Maturity Between 1956 and 1977 ..... | 95  |
| <i>U. Dash and T. Maguire</i>  |     |
| An Empirical Analysis of Loans by School Libraries .....   | 104 |
| <i>R. Houle and C. Montmarquette</i>   |     |
| Serving on the School Board:<br>A Political Apprenticeship for Higher Office .....   | 115 |
| <i>N. Robinson and C. Stacey</i>   |     |
| The Relationship Between Simultaneous-Successive Processing<br>and Academic Achievement .....                                | 126 |
| <i>F. M. Merritt and R. S. McCallum</i>  |     |

ESSAY REVIEWS

|  |     |
|--|-----|
| Logo and the Development of General Problem-Solving Skills .....                       | 133 |
| <i>L. R. Krasnor and J. O. Mitterer</i>  |     |
| Self-fulfilling Prophecies in the Classroom:<br>Dead End or Promising Beginning? ..... | 145 |
| <i>V. S. Smead</i>   |     |

FACULTY OF EDUCATION  
*The University of Alberta*



DONALD E. ALLISON

*University of British Columbia*

## The Effect of Homogeneous vs. Heterogeneous Matching-Item Format on Test Performance and Reliability

*Two forms of a general science matching-item test were prepared. Both forms consisted of three matching exercises. Each of the exercises in Form A contained homogeneous items devoted to a single topic while the exercises in Form B contained a heterogeneous arrangement of items drawn at random from the topics in Form A. When the test was administered to 316 sixth grade students, analysis of the data revealed that, while there was no significant difference between the reliability of the two forms, scores on the heterogeneous form of the test were significantly higher than those on the homogeneous form and that this difference was independent of any interaction with the examinee's sex or intelligence.*

Although they do not cite empirical evidence to support their contention, authors of measurement books continue to recommend that matching exercises should be homogeneous (Brown, 1983; Hills, 1981; Hopkins & Antes, 1978; Lien, 1980; Mehrens & Lehmann, 1975; Thorndike & Hagen, 1977). Homogeneity, in this case, means stimuli and responses that are restricted to single concepts, classifications, or areas, so that a homogeneous matching-item test would not likely include persons, places, and things in a single exercise. For the most part, these authors maintain that homogeneous matching exercises require students to demonstrate their knowledge of the subject content by making careful discriminations in selecting the correct responses. On the other hand, exercises which are not homogeneous provide test-wise students with clues which allow them to eliminate responses which have no logical connection to the stimulus, thus enabling them to select the correct responses from a much shorter and easier list.

This study was designed to make an empirical comparison between two formats, homogeneous and heterogeneous, of a matching-item test. Since any difference be-

---

Donald Allison (PhD, Southern California) is an Associate Professor and Coordinator of the Measurement, Evaluation and Research Methodology area of the Department of Educational Psychology and Special Education, The University of British Columbia. His current research and publications are focused on classroom testing and test construction.



tween the difficulty and reliability of the homogeneous and heterogeneous forms of the test might be due to an advantage that a heterogeneous matching exercise might give to more able and/or test-wise students, less experienced and sophisticated elementary school students were chosen as subjects for this investigation.

It was hypothesized that there is no significant difference in test performance due to (1) matching-item format or (2) any possible interaction between matching-item format, intelligence and sex, or any combination of these three variables, and (3) that there is no significant difference in test reliability due to matching-item format.

### *Method*

#### *The Instrument*

Two forms of a sixth grade general science matching-item test were constructed. Form A consisted of three homogeneous, 10-item, 11-option, matching exercises. The first exercise was concerned with the classification of animals, the second with the solar system, and the third with earth science. Each of the exercises was placed on a separate page with the options arranged in alphabetical order. Form B also consisted of three, 10-item, 11-option, matching exercises, but the items and their matching options were drawn at random from Form A. Form B exercises, then, were heterogeneous, containing items from all three topics. As with Form A, each exercise was placed on a separate page with the options in alphabetical order. The maximum possible score for both Forms A and B was 30.

#### *Procedure*

The subjects were 316 sixth grade students enrolled in 15 classes in a suburban school district. Forms A and B were distributed alternately within each classroom, so that when all the students had been tested, 156 had completed Form A and 160 Form B.

From intelligence test scores obtained from school records, the mean IQ of these students was found to be 112.37 and the median 113.63. When students were classified as high- or low-IQ by dividing the intelligence test scores at the median, there were 159 high-IQ students and 157 low-IQ students.

### *Results*

The means and standard deviations of the matching-item test scores, when students were classified by intelligence (IQ), sex, and test format, are presented in Table 1. This table shows that the mean score for the total group of 316 students was 18.87 with a standard deviation of 6.06.

The significance of these data was assessed by means of a 2 (IQ) X 2 (sex) X 2 (format) analysis of variance. This analysis is summarized in Table 2.

The results show that the main effect due to format was significant,  $p < .01$ , causing rejection of the first hypothesis that there would be no difference in test performance due to matching-item format. Table 1 reveals that scores were higher on the heterogeneous format (Form B) than on the homogeneous format (Form A).

No hypotheses related to the other two main effects had been stated. The significant effects ( $p < .01$ ) due to intelligence and sex indicated that the high-IQ students did better on the matching-item test than the low-IQ students and that the science content of the test was easier for the boys than for the girls.

TABLE 1  
MEANS AND STANDARD DEVIATIONS OF MATCHING-ITEM  
TEST SCORES: STUDENTS CLASSIFIED BY IQ, SEX, AND FORMAT

| Variable | Level | <u>n</u> | Mean  | <u>SD</u> |
|----------|-------|----------|-------|-----------|
| IQ       | High  | 159      | 21.81 | 5.04      |
|          | Low   | 157      | 15.89 | 5.55      |
| Sex      | Boys  | 157      | 19.87 | 5.88      |
|          | Girls | 159      | 17.88 | 6.10      |
| Form     | A     | 156      | 16.67 | 5.45      |
|          | B     | 160      | 21.01 | 5.87      |
| Combined |       | 316      | 18.87 | 6.06      |

TABLE 2  
SUMMARY OF THE ANALYSIS OF VARIANCE OF MATCHING-ITEM  
TEST SCORES BY INTELLIGENCE, SEX, AND FORMAT

| Source of Variation | <u>df</u> | Mean Square | <u>F</u> |
|---------------------|-----------|-------------|----------|
| IQ                  | 1         | 2656.973    | 116.361* |
| Sex                 | 1         | 270.862     | 11.862*  |
| Format              | 1         | 1505.178    | 65.918*  |
| IQ X Sex            | 1         | 24.100      | 1.055    |
| IQ X Format         | 1         | 7.059       | 0.309    |
| Sex X Format        | 1         | 0.086       | 0.004    |
| IQ X Sex X Format   | 1         | 10.966      | 0.480    |
| Within Cells        | 308       | 22.834      |          |
| Total               | 315       |             |          |

\*p < .01

As predicted in the second hypothesis, Table 2 also shows that none of the interactions between IQ, sex, and format were significant,  $p < .05$ .

Kuder-Richardson formula 20 reliabilities were calculated as .777 and .845 for Forms A and B of the test. The difference between these two coefficients, when assessed by the procedure suggested by Feldt (1969) was not found to be significant,  $W(155,159) = 1.439$ ,  $p > .05$ , thus supporting the third hypothesis, that there is no difference in test reliability due to matching-item format.

### Discussion

These results show that even when the subject matter content is identical, the format of a matching-item test influences test performance, and this effect is independent of any interaction with either the sex or intelligence of the examinees. While there was no significant difference in the reliability, scores on the heterogeneous form of this test were significantly higher than those on the homogeneous form, leading to a conclusion that the difference was due to the matching-item format rather than knowledge of the content.

Since scores on an achievement test should be a function of content mastery rather than item format, it can also be concluded that measurement specialists are correct in emphasizing the importance of homogeneity in the design and construction of matching-item exercises.

### References

- Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). New York: Holt, Rinehart and Winston.
- Feldt, L. S. (1969). A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. *Psychometrika*, 34, 363-373.
- Hills, J. R. (1981). *Measurement and evaluation in the classroom* (2nd ed.). Columbus, OH: Charles E. Merrill.
- Hopkins, C. D., & Antes, R. L. (1978). *Classroom measurement and evaluation*. Itasca, IL: F. E. Peacock.
- Lien, A. J. (1980). *Measurement and evaluation of learning* (4th ed.). Dubuque, IA: Wm. C. Brown.
- Mehrens, W. A., & Lehmann, I. J. (1975). *Measurement and evaluation in education and psychology* (2nd ed.). New York: Holt, Rinehart and Winston.
- Thorndike, R. L., & Hagen, E. P. (1977). *Measurement and evaluation in psychology and education* (4th ed.). New York: John Wiley and Sons.



UDAYA DASH

*Utkal University*

and

THOMAS MAGUIRE

*University of Alberta*

## A Detailed Analysis of Group Differences on the California Short-Form Test of Mental Maturity Between 1956 and 1977

*The California Short Form Test of Mental Maturity was administered to 3,443 third grade students in 1956, and to 4,378 third grade students in 1977. Differences in factorial structure were examined, as were differences in ability level between the two groups on those factors which appear to be measuring consistent abilities. Of the 98 items, 64 appeared to behave with factorial consistency in the two administrations. Consistent items were related to seven components including two kinds of Sensing Right and Left, Spatial Relationships, Similarities, Numerical Reasoning, Vocabulary and Verbal Mediation. Differences in ability level favored the 1977 group on Sensing Right and Left II, Numerical Reasoning and Verbal Mediation. The 1956 group had a higher mean than the 1977 group on Vocabulary. Results were interpreted in terms of Jensen's ability levels and in terms of specific changes in curriculum and children's television.*

There have been a number of reports showing that educational test scores have declined over the past few decades (Harnischfeger & Wiley, 1976).<sup>1</sup> Closer examination reveals that the phenomenon is more complex than was first thought. During the 1940s and through to the mid-'60s, scores seemed to increase fairly steadily, with declines occurring in the past two decades. The declines have tended to be greater in the upper grades than in the elementary grades and there seemed to be differential effects related to student gender.

---

Udaya Dash is a lecturer in the Department of Psychology, Utkal University, India. He is interested in measurement, applied statistics, and cognitive development.

Thomas Maguire is a professor in the Department of Educational Psychology, University of Alberta. His interests are educational measurement, applied statistics, and research methodology.

Reports on score decline at the elementary school level are not unequivocal. As Harnischfeger and Wiley note, there has been little or no decline in Iowa Tests of Basic Skill scores at the third grade, and there has actually been a modest increase for the Comprehensive Tests of Basic Skills at grades 2, 3, and 4. On the other hand, National Assessment of Educational Progress scores have shown decreases in science for 9- and 13-year-old students during the early seventies.

In Canada, studies over 40 years (Hedges, 1977) and over 21 years (Clarke, Nyberg, & Worth, 1977) have shown the same mixture of results, with declines being noted among older children and no differences or modest increases occurring with younger children.

Although less well-studied than achievement, school-related ability measures provide a similar pattern. In high schools, declining scores have been noted in the Scholastic Aptitude Test, the American College Testing Program, the Preliminary Scholastic Aptitude Test, and the Minnesota Scholastic Aptitude Test. At grade 3, Nyberg and Blackmore (1981) report an increase of about one-half of a standard deviation in raw scores on the California Short Form Test of Mental Maturity.

There have been many speculations concerning the cause for changes in aptitude and achievement scores, amongst which the cumulative effect of TV, retention of proportionately more students in high school, and changes in curriculum have been most prominent. Unfortunately, interpretations have been made difficult by the omnibus nature of the tests used. It seems likely that investigation into the nature of aptitude and achievement differences would be more profitable if attention were paid to differential changes in specific abilities, or specific achievements at the subtest level, or even at the item level. For example, one might expect TV to influence negatively some areas such as reading comprehension, but science knowledge or knowledge of current affairs might be changed positively. Changes in curriculum ought to manifest themselves in a fairly specific fashion, whereas greater school retention rates may lead to more general changes. Thus to begin to understand the nature of the phenomenon, it is necessary to look for differences in specific areas.

### *Objectives*

The purpose of the present study was to examine specific changes in educationally related abilities for two groups of third grade students measured in 1956 and 1977 using the California Short Form Test of Mental Maturity (CSTM)—Primary S. The samples were made up of all third grade students in the Edmonton Public Schools who were present on the testing days in those two years. More specifically the objectives were:

1. To examine differences in factorial structure of item responses between 1977 and students in 1956 on the CSTM.
2. To examine differences in ability level between the two groups on those factors which appear to be measuring consistent abilities.

### *Method*

#### *Sample*

The sample consisted of 7,821 Edmonton Public School Board Grade 3 students: 3,443 children were tested in 1956 and 4,378 children were tested in 1977. All children were tested on a large battery of cognitive tests for the project described by Clarke, Nyberg and Worth (1977). The present paper is an extended analysis of some of the data used by Clarke et al. (1977), by Nyberg and Blackmore (1981), and by Blackmore (1980).

Test

The California Short Form Test of Mental Maturity (CSTM), now out of print, was a part of the larger parent test called the California Test of Mental Maturity. The CSTM consists of 98 items grouped into 7 subtests (Sensing Left and Right, Manipulation of Areas, Similarities, Inference, Number Series, Numerical Quantity, and Verbal Concepts) which are used to obtain a general intellectual maturity measure in a one-period group test. The test, its reliability, validity, administration, and scoring procedures are reported in the test manual (Sullivan, Clark, & Tiegs, 1953).

According to Clarke et al. (1977), teachers, supervisors in elementary education from the school system, and experts from the University of Alberta were asked to identify items which would be inappropriate for the 1977 class. In the case of the CSTM no such item was reported. The CSTM makes provision for calculating an IQ score based on the total score for the test. Although time limits are imposed on the students, the test is basically designed as a power test. In Table 1, a detailed description of each of the subtests is given. The subtests are grouped according to what Sullivan et al. call, “major factors involved in intelligence or mental capacity.” These are: spatial relations, logical reasoning, numerical reasoning, and verbal concepts.

TABLE 1  
DESCRIPTION OF THE CALIFORNIA SHORT-FORM  
TEST OF MENTAL MATURITY

**Spatial Relationships—Orientation in space and the use of spatial relationships**

*Test 1: Sensing Right and Left*

The first part of the test consists of 6 pictures of girls and boys and the students are instructed to mark particular appendages such as, “Put a mark on the girl’s left arm.” The second part of the test consists of four double pictures of hands or feet in various positions and the students are told to “Put a mark on each right hand or foot.”

*Test 2: Manipulation of Areas*

The test consists of 12 picture items intended to measure the students’ ability to use spatial imagery to manipulate different forms. Each item consists of a row of four drawings. The student is told that the “first drawing is among the other drawings but is turned around or turned over. Find it and put a mark on it.”

**Logical Reasoning**

*Test 3: Similarities*

This test consists of a total of 12 picture situations. The first two pictures are alike in some ways (e.g., monkey, horse). The pupil determines the nature of this likeness and finds another picture among the other three (tea pot, broom, mouse,) which is similar to the first two.

*Test 4: Inference*

The test is made up of 12 picture situations. The examiner reads a few short sentences and then asks the students to make an inference. For example, in one situation there is a picture of two boys holding fish. The examiner says, “Look at the two boys. Bill caught more fish than Ned. Put a mark on Bill. He is the boy with the most fish.”

**Numerical Reasoning**

*Test 5: Number Series*

There are 12 picture situations which require the pupil to understand lightest, heaviest, most, etc., and number series. Six of the situations consist of pictures (e.g., bicycle, plane, car)



and the student is required to choose the one that falls at the extreme on some characteristic (e.g., put a mark on the thing that can go the fastest). In the other six situations there are five boxes with different numbers of circles marked on them to form a number series. The student is told, "One box is wrong in each row. Put a mark on the box that is wrong."

*Test 6: Numerical Quantity*

There are 12 picture problems. The student is given a numerical situation and is asked a question. (e.g., "Look at the first box in this row. It has three blocks. If you take one of the blocks away which box will it look like?" Three boxes follow with 2, 5 and 4 blocks in them.)

*Test 7: Verbal Concepts*

This test consists of 28 items each containing three pictures. The student is given a single word or concept and required to choose the picture that matches the word or concept. For example, the three pictures are: children going upstairs, a merry-go-round, children going downstairs. The examiner says, "Put a mark on those descending."

*Data Analysis*

The two inter-item correlation matrices, one for the 1956 and the other for the 1977 sample, were analyzed using a principal components analysis. On the basis of the scree test, seven components were retained and rotated to the varimax criterion. The varimax loadings of the 1956 sample were rotated to match the loadings obtained for the 1977 sample using an orthogonal Procrustes solution. The error matrix was used for identifying the items that contributed to differences in factor structure. Any item having a difference of .15 in its loadings on the same component for the two groups was removed. In addition, any item which had differential loadings of .10 or more on *two or more* components was also removed.

Although the sampling distribution of factor loadings is not well explicated, it was felt that with over 300 subjects these criterion values would not only be significant, but important as well. Accordingly, 34 out of the 98 items in the CSTM were removed. The intercorrelation matrix of the remaining 64 items for both groups combined was analyzed and rotated to the varimax criterion. Component scores were calculated for all the subjects. Analysis of variance was performed on the seven component scores between 1956 and 1977 students to compare the achievement differences of both groups, on those parts of the test having a common structure.

*Results and Discussion*

*Differences in Component Structure*

Thirty-four items were found to contribute to differences in component structure. In the case of 27, the difference between a rotated component loading for 1956 and a 1977 component loading exceeded .15. (The largest discrepancy was .23.) For 9 of the 27, this occurred on the defining component. Seven of the 34 items were excluded because of two or more discrepancies between .10 and .14. The source of the items which did not fit is shown in Table 2. From that table, it can be seen that there seems to be a disproportionately small number from Right Hands and Feet, Manipulation of Areas, and Logical Inference. In spite of these differences, viewed from an item salience perspective, the fit was remarkably good. On all but 14 items the highest component loading for the 1977 solution agreed with that of the rotated 1956 version.

Factor Structure of the CSTM

Seven components emerged clearly and consistently when separate analyses were performed on the 1956 and 1977 groups, and also when the combined data of both groups on the 64-item portion of the test were analyzed. The first component was defined by items that required subjects to identify the left and right hand or foot of single pictures of boys or girls. This factor was named Sensing Left and Right—I. The second component, which was named Sensing Left and Right—II, was defined by items requiring relatively more complex discriminations between left and right than in the first. They dealt with double picture situations at a relatively more difficult level of operation. The third component was labelled Spatial Relationship, since the items with high loadings required subjects to identify a rotated or an inverted image of the target pattern in the midst of a number of distracting patterns. These items appear to measure subjects' ability to use spatial imagery in manipulating patterns of different forms and in many different posi-

TABLE 2  
SOURCES OF ITEMS THAT CONTRIBUTED  
TO DIFFERENCES IN FACTOR STRUCTURE

| Subtest               | Items<br>Contained<br>in Subtest | Items<br>Contributing<br>to Difference | Proportion |
|-----------------------|----------------------------------|--|------------|
| Sensing Right & Left  |                                  |  |            |
| Body Parts            | 1 to 6                           | 3,5,6                                  | .5         |
| Right Hand or Foot    | 7 to 12                          |  | 0          |
| Manipulation of Areas | 1 to 12                          | 6,12                                   | .17        |
| Similarities          | 1 to 12                          | 2,6,8,10                               | .33        |
| Inference             | 1 to 12                          | 7,10                                   | .17        |
| Number Series         |                                  |  |            |
| Choose Extreme        | 1 to 6                           | 3,5,6                                  | .5         |
| Number Series         | 7 to 12                          | 11,12                                  | .33        |
| Numerical Quality     | 1 to 12                          | 1,2,3,7,8,9                            | .5         |
| Verbal Concepts       | 1 to 28                          | 1,2,3,5,6,<br>10,12,14,16,<br>19,21,24 | .43        |
| TOTAL                 | 98                               | 34                                     | .35        |

tions. The fourth component, named Similarities, was defined by items which asked subjects to examine the similarities between two stimulus pictures, and then to identify a picture among the three response pictures that is similar to the stimulus. The fifth component, Numerical Reasoning, consisted of items measuring pupils' ability to make inferences with special reference to quantitative situations and problems. The sixth component, Vocabulary, was defined by items which require the subjects to find pictures of certain objects or animals among a number of other pictures. The items that loaded on the seventh component, Verbal Mediation, test pupils' ability to identify pictures that portray the verbal descriptions specified by the items. (For example, two of the items under this component required subjects to find "something delicious," and find "those descending.") It is interesting that subtest 4 (Inference) did not belong to the space occupied by other items in the test. Indeed when the correlation matrix was examined, none of the 45 correlations between Inference items exceeded .1, and only 5 of the 540 correlations between Inference items and non-Inference items exceeded .1, the largest being .13.

Table 3 presents the clustering of the 64 items on the seven components. The emergence of these seven components in a sample size much larger than the sample (N=700) on which the original structure of the CSTM was based offers support for the validity of the original structure of the CSTM, except of course for the Inference dimension.

TABLE 3  
CLUSTERING OF THE 64 ITEMS ON SEVEN COMPONENTS<sup>a</sup>

| Components                    | Subtest No. | Item No.                        |
|-------------------------------|-------------|---------------------------------|
| I. Sensing left and right-I   | 1           | 1,2,4                           |
| II. Sensing left and right-II | 1           | 7,8,9,10                        |
| III. Spatial relationship     | 2           | 1,2,3,4,5,7,8,9,10,11           |
| IV. Similarities              | 3           | 1,3,4,5,7,9,11,12               |
| V. Numerical reasoning        | 5<br>6      | 1,2,4,5,6,7,8<br>4,5,6,10,11,12 |
| VI. Vocabulary                | 7           | 4,7,8,9,11,13,15,18             |
| VII. Verbal mediation         | 7           | 17,20,22,23,25,26,<br>27,28     |

<sup>a</sup> The item Nos. 1,2,3,4,5,6,8,9,11,12 of Subtest No. 4 (Inference) did not load on any component.



Ability Differences Between 1956 and 1977 Groups

Table 4 presents the mean student abilities of 1956 and 1977 Edmonton Grade 3 children based on component scores calculated from the 64 consistent items of the CSTM as well as on the component scores derived from the entire test.

TABLE 4  
MEAN STUDENT ABILITIES OF 1956 AND 1977 EDMONTON  
GRADE 3 CHILDREN AND F VALUES ON SEVEN COMPONENTS

| Factors                   | 1956  | 1977  | F        |
|---------------------------|-------|-------|----------|
| Sensing left and right-I  | .037  | -.029 | 7.04     |
| Sensing left and right-II | -.454 | .357  | 1524.90* |
| Spatial relationship      | -.031 | .026  | 7.11     |
| Similarities              | -.023 | .020  | 3.76     |
| Numerical reasoning       | -.138 | .109  | 112.13*  |
| Vocabulary                | .111  | -.084 | 72.09*   |
| Verbal mediation          | -.280 | .221  | 521.32*  |

Based on 64 items

\* Significant <.001

Given the large number of students in the sample, it was decided to use a significance level of .001. This difference between means corresponds to about a tenth of a component score standard deviation. (Significance at the .05 level corresponds to a difference of about one/twenty-fifth of a standard deviation.) By using the higher critical difference, it was hoped that trivial differences could be avoided and as well the overall significance level would be kept below .01.

As may be seen from Table 4, the 1977 group performed at a higher level than the 1956 group on the components labelled Sensing Right and Left II, Numerical Reasoning, and Verbal Mediation. In Vocabulary the 1956 group had a higher mean than the 1977 group.

Several speculations can be made as to the reasons for these differences. Vocabulary as a cognitive skill does not emphasize the processing requirements that

Sensing Left and Right II, Numerical Reasoning, and Verbal Mediation do. In contrast to Vocabulary, the processing requirements of the other three components correspond to Jensen's (1970) Level II ability in the sense that the original input needs to be manipulated, evaluated, and transformed to produce an effective output. On the items defining Vocabulary, the subjects are merely required to point to one of three pictures. The processing requirements are considerably less in Vocabulary for most people.

If curriculum and instruction in 1956 placed more emphasis on knowledge than in 1977, and if a gradual shift toward emphasizing understanding, analysis, and evaluation has taken place since, then the present results would be consistent. The largest difference occurred on Sensing Right and Left II which requires the student to be able to distinguish right and left hands and feet, and to identify them under different orientations. Introduction of manipulative aids to primary school arithmetic may partially account for the observed differences. The kind of skills taught on children's television would also contribute to this difference.

Both Numerical Reasoning and Verbal Mediation involve aural memory and reasoning which are also skills promoted by children's television. The items for Logical Inference (Test 4) which did not contribute to test structure may have exceeded the capability of the children at this level to retain the information in memory and then to carry out the complex analysis required. Even so, the children may have been able to combine intuition with partial information to solve the problems. Whatever the cause, on eleven of the twelve items, the 1977 group had higher means, which supports the contention that aural memory and reasoning skills are slightly higher in the 1977 group.

In their 1981 article, Nyberg and Blackmore note a difference in CSTM mean from 1956 to 1977 of about two-thirds of a standard deviation. Blackmore's (1980) latent trait analysis of the same data showed a difference of about one-half a standard deviation on the ability scale after 25 aberrant items have been removed. In both cases the results favored the 1977 group. The results of the present study suggest that the differences across abilities are not constant and indeed the order is not even consistent. The 1977 students seem to have superior level II processing skills, whereas the 1956 students are favored on vocabulary as measured by picture recognition.

Because the results differentially favor one group in some cases and the other group in other cases, it seems unlikely that global influences of population migration or teacher recruitment are operating. Such influence would seem more likely to produce general differences over many areas favoring a single group. More plausible influences would seem to be specific changes in curriculum and television viewing habits. It is impossible to document all the changes that took place in the 21 years, but closer examination of the data narrows the search. Detailed analysis of the remaining tests in the Clarke, Nyberg, and Worth (1977) battery is being undertaken with the hope of specifying the changes in skills more clearly.

#### Notes

1. Harnischfeger and Wiley provide an excellent summary of literature in the area. Unless otherwise stated, research cited comes from that review.

References

- Blackmore, D. E. (1980). *A latent trait study of item biachievement differences*. Unpublished doctoral dissertation, University of Alberta, Edmonton, Alberta.
- Clarke, S. C. T., Nyberg, V., & Worth, W. H. (1977). *General report on Edmonton Grade III achievement: 1956-1977 comparisons*. Edmonton, Canada: Alberta Education.
- Harnischfeger, A., & Wiley, D. (1976). Achievement test scores drop. So what? *Educational Researcher*, 5, 5-12.
- Hedges, H. G. (1977). *Achievement in basic skills: A longitudinal evaluation of pupil achievement in language arts and mathematics*. Toronto: Ontario Institute for Studies in Education.
- Jensen, A. R. (1970). Hierarchical theories of mental ability. In W. B. Dockerell (Ed.), *On intelligence*. Toronto: Ontario Institute for Studies in Education.
- Nyberg, V., & Blackmore, D. E. (1981). A longitudinal study of Grade III achievement in Edmonton Public Schools. *Alberta Journal of Educational Research*, 27, 154-159.
- Sullivan, E. T., Clark, W. W., & Tiegs, E. W. *California shortform test of mental maturity, Primary grades 1, 2, 3, 1953 S-Form*. Los Angeles: California Test Bureau.



RACHEL HOULE

and

CLAUDE MONTMARQUETTE

*Université de Montréal*

## An Empirical Analysis of Loans by School Libraries

*This empirical study measures the influence of selected factors on loans of printed materials made by a sample of Québec primary and secondary school libraries. A production function approach is chosen with library loans as the output, and library resources and many policy and interaction variables as inputs. The analysis relies on probit and linear regression techniques for the estimation of parameters after a distinction is made between school libraries that make loans and those which do not. The study yields information to help shape policies about the number of loans by school libraries.*

Reading is not only one of the basic skills to be acquired but, as well, is considered by both parents and professional educators to be a pedagogical instrument to stimulate the intellectual development of children. School libraries provide both reading materials and a supportive environment and appear, therefore, to constitute a key element in promoting the importance and enjoyment of reading. However, the efficiency of school libraries in developing reading habits is conditional on certain factors among which is the allocation of the meager \$5 and \$20 per student spent annually, from 1972 to 1979, for Québec primary and secondary school libraries respectively.

---

Rachel Houle has specialized in regional economics and economics of education at the University of Pennsylvania where she obtained her PhD in 1979. Currently Assistant Professor of Economics at Université de Montréal, Miss Houle has produced several scientific articles.

Claude Montmarquette has specialized in public finance and econometrics at the University of Chicago where he obtained his PhD in 1973. Currently Professor of Economics and Director of the Centre de Recherche en Développement Economique (C.R.D.E.), Université de Montréal, Mr. Montmarquette has published several scientific articles.

Library use is a multidimensional concept. School library materials can be consulted on the premises or borrowed; the users can be students, teachers, parents, or other people; there are some inter-library loans; and some libraries also have audio-visual materials. Assuming that book borrowing influences reading habits, the focus of this study is on the loans of printed materials to individuals. Unfortunately, the use of library materials on the premises had to be excluded from the analysis because of the lack of reliable data.

We estimate the role of selected factors on loan policies of school libraries at the elementary and secondary levels by providing answers to the following questions. What part do the quantity and quality of books in the library, the size and composition of personnel, the organizational methods, the physical facilities, and the users' characteristics play in determining whether or not a school library makes loans? For those school libraries making loans, what are the main determinants of the number of loans per pupil made over a school year?

The quantitative analysis performed in the study links the output of school libraries (loans) to their inputs (e.g., number of books or personnel) in a production function framework. The analysis includes numerous explanatory interaction variables and relies on particularly suitable statistical models, namely the probit and multiple regression techniques.

The analysis should yield interesting information to help shape policies designed to increase the number of loans made by school libraries. In a period of severe financial constraint currently affecting education in Québec and elsewhere in Canada, the question of the best use of all available resources is of major importance.

### *A Model of the Determinants of School Library Loan Services*

Contrary to the recent public library literature, the literature on the determinants of school library services at the elementary and secondary levels is very scarce. The few studies on the subject are more descriptive than causal and the discussion is generally qualitatively oriented.<sup>1</sup> The following services associated with school libraries are recognized in this descriptive literature:

1. technical services: selection, acquisition, classification and conservation of materials, and
2. services to users such as information, reference and loans.

To include in the model all services performed by school libraries would require the analysis to be conducted in a multi-output production function framework. Unfortunately, the number of loans is the only output for which reliable quantitative information exists at the present time for a large number of elementary and secondary schools, so that such a framework cannot be employed.

The output variable, the level of which constitutes the explanatory focus, takes two forms in the analysis. The first form of the output variable is a dichotomous variable, that is, a variable taking only two values—whether or not a school library makes loans; the second is the number of loans per pupil made over a school year provided that a school library makes loans.

In a causal model, two sets of explanatory variables are generally distinguished: first, *the policy variables* connected with the human and physical resources and the organization and, second, *the standardizing variables*, which in this case are the users' and the school's characteristics. This second set of variables lies outside the school library operating policies.

TABLE 1

MEAN VALUES AND STANDARD ERRORS  
OF SELECTED VARIABLES

| Variables  | Elementary schools |                             | Secondary schools |                 |
|--|--------------------|-----------------------------|-------------------|-----------------|
|  | full sample        | with loans                  | full sample       | with loans      |
| Number of schools                                  | 877                | 417                         | 350               | 283             |
| Number of annual loans (per pupil)                 |                    | 38.0<br>(30.5) <sup>a</sup> |                   | 17.18<br>(16.1) |
| Number of Books (per pupil)                        | 16.12<br>(8.35)    | 16.5<br>(9.03)              | 23.1<br>(20.3)    | 22.7<br>(19.6)  |
| Number of periodicals (per pupil)                  | .06<br>(.24)       | .07<br>(.29)                | .1<br>(.6)        | .11<br>(.45)    |
| New acquisitions (% of total books)                | .08<br>(.18)       | .10<br>(.22)                | .15<br>(1.6)      | .06<br>(.08)    |
| Weedings (% of total books)                        | .02<br>(.06)       | .02<br>(.06)                | .03<br>(.09)      | .03<br>(.09)    |
| Library personnel per 100 pupils:                  |                    |                             |                   |                 |
| professionals (full-time)                          | -b                 | -b                          | .05               | .06             |
| professionals (part-time)                          | .1                 | .14                         | .06               | .07             |
| technicians (full-time)                            | -b                 | .01                         | .06               | .06             |
| technicians (part-time)                            | .06                | .08                         | .05               | .05             |
| support staff (full-time)                          | .01                | .01                         | .07               | .06             |
| support staff (part-time)                          | .19                | .28                         | .09               | .08             |
| Number of operating hours (per week)               | 19.7<br>(10.8)     | 22.6<br>(9.2)               | 30.31<br>(12.4)   | 32.2<br>(10.9)  |
| Library area (in square meters per pupil)          | .33<br>(.48)       | .36<br>(.43)                | .62<br>(.86)      | .61<br>(.85)    |
| Number of carrels (per pupil)                      | .0032<br>(.04)     | .0045<br>(.05)              | .01<br>(.02)      | .01<br>(.02)    |
| Number of chairs (per pupil)                       | .1<br>(.1)         | .12<br>(.09)                | .12<br>(.11)      | .12<br>(.1)     |
| Number of schools in socioeconomic areas which are |                    |                             |                   |                 |
| strongly underprivileged                           | 176                | 71                          | 49                | 43              |
| underprivileged                                    | 324                | 146                         | 133               | 102             |
| privileged   | 377                | 200                         | 168               | 138             |
| Number of schools located in towns                 |                    |                             |                   |                 |
| of less than 10 000 inhabitants                    | 492                | 221                         | 161               | 131             |
| of 10 000 to 25 000 inhabitants                    | 122                | 62                          | 56                | 42              |
| of 25 000 to 100 000 inhabitants                   | 192                | 100                         | 87                | 73              |
| 100 000 inhabitants or more                        | 71                 | 34                          | 46                | 37              |
| Number of teachers in school                       | 15.2<br>(7.0)      | 16.1<br>(7.7)               | 56.3<br>(45.2)    | 62.6<br>(45.9)  |
| Number of private schools                          | 9                  | 5                           | 56                | 43              |

<sup>a</sup> Standard Error shown in parentheses

<sup>b</sup> Negligible value



TABLE 2  
TOTAL ELASTICITIES

| Variables                                       | Elementary school                             |                       | Secondary school                              |                       |
|---|---|-----------------------|---|-----------------------|
|   | probabil-<br>ities of<br>loans <sup>a,c</sup> | number<br>of<br>loans | probabil-<br>ities of<br>loans <sup>b,c</sup> | number<br>of<br>loans |
| <u>Size and quality of the stock:</u>           |   |                       |   |                       |
| Number of books (per pupil)                     | -.042   | .454<br>(.115)        | .044  | .437<br>(.157)        |
| Number of periodicals (per pupil)               | .023  | .148<br>(.027)        | .003  | -.076<br>(.067)       |
| New acquisitions (% of total books)             | .194  | .045<br>(.074)        | -.81E-07                                      | .026<br>(.087)        |
| Weedings (% of total books)                     | .325  | .059<br>(.065)        | .176  | .041<br>(.038)        |
| <u>Library personnel:</u>                       |   |                       |   |                       |
| Professionals (full-time)                       | -.022   | -.01<br>(.033)        | .034  | .016<br>(.05)         |
| Professionals (part-time)                       | .013  | -.023<br>(.007)       | -.004   | .015<br>(.038)        |
| Technicians (full-time)                         | .003  | .028<br>(.03)         | .012  | -.003<br>(.045)       |
| Technicians (part-time)                         | .086  | .039<br>(.023)        | .165  | .078<br>(.032)        |
| Support staff (full-time)                       | .359  | .079<br>(.053)        | -.017   | -.032<br>(.033)       |
| Support staff (part-time)                       | .034  | -.003<br>(.017)       | .002  | .058<br>(.023)        |
| <u>Library facilities:</u>                      |   |                       |   |                       |
| Number of operating hours (per week)            | .299  | .175<br>(.11)         | .02   | .832<br>(.307)        |
| Library area (in sq. m. per pupil)              | .052  | -.127<br>(.071)       | -.009   | .16<br>(.096)         |
| Number of booths (per pupil)                    | -.009   | -.027<br>(.014)       | .005  | -.059<br>(.042)       |
| Number of chairs (per pupil)                    | .184  | .132<br>(.079)        | -.019   | .146<br>(.111)        |
| <u>Others:</u>                                  |   |                       |   |                       |
| Percentage of 1st to 3rd graders                | .169  | .114<br>(.091)        | -   | -                     |
| Percentage of students in lower grades          | -   | -                     | .062  | .187                  |
| Percentage of students in professional programs | -   | -                     | .005  | -.022<br>(.048)       |
| Number of teachers                              | .198  | .055<br>(.086)        | .047  | -.06<br>(.107)        |

<sup>a</sup> Computed at mean values of variables and coefficients of regression results reported in tables of the appendix.

<sup>b</sup> Computed at mean value of variables minus two standard errors and coefficients of regression results. Owing to a negative value, we substitute the computed value by 1E-06, for 6 variables.

<sup>c</sup> Estimated standard errors are not reported for the probability of loans elasticities (elementary schools): all were found relatively large but remain questionable due to non linearity and their approximation from a Taylor series expansion.

Source: Listings L11DP02, L11DPY4, L11DPYF, L11DPN1, L11DPOE, L11DPF1, L11DP8Y, L11DPS4, L11DPGY, L11DP00, L11DPXU, L11DPMG.

### *Policy Variables<sup>2</sup>*

1. *The size and quality of the stock of printed materials* is measured by the number of books and periodicals per pupil, new acquisitions as a percentage of the total number of books, and the previous year's weeding as a percentage of the total number of books. These last measures are used as *proxies* for the quality of books; the hypothesis employed is that library personnel acquire books and dispose of others to improve the stock. Eliminating outdated or dilapidated books should increase the access to the remaining books. The size and quality of the stock are expected to influence directly and positively the probability that a library makes loans and the number of loans made.

2. *The library personnel* are classified as professionals, technicians or support staff working full-time or part-time. Professionals include professional librarians, educational advisors or specialists in learning materials and methods as well as teachers participating in their school library's operation. Technicians are employees who have received some training in library management and who work under the supervision of a professional. Library employees are also expected to influence directly and positively the output variables. Furthermore, as additional personnel should improve the efficiency of the management of the stock of books, interaction variables between the quantity and quality of book variables and the library personnel variables should also have a positive effect on the output variables. To illustrate the concept of interaction effect, let us consider the interaction variable defined as "weeding x full-time professionals." The use of this variable implies that the influence of the variable "weeding" on the probability of loans and on the number of loans also depends on the number of professionals working full-time in the school library.<sup>3</sup>

3. *The library organization and facilities* are taken into account by including the number of operating hours per week, the total library area, and the number of chairs and carrels available per pupil. The first two variables should influence the number of loans positively but the last two variables may simply induce pupils to work or read on the premises rather than borrow the materials for use outside the library.

A stronger negative effect is expected if the school encourages the formation of separate collections of books in classrooms, in addition to the existence of a central library. Finally, as for the previous class of factors, some interaction variables between library organization and facilities and (1) library personnel, (2) number of books, are included in the causal model.

### *Standardizing Variables*

*Users' characteristics* are considered as playing an important role in determining the use of school libraries.<sup>4</sup> Four characteristics of students are selected as particularly relevant:

1. *The age of students* is expected to influence loans positively, since older students are expected to read more both for classwork and leisure. However some empirical studies (Houle, 1980) on reading habits suggest the opposite hypothesis. The influence of age is captured through variables measuring the fraction of students in the lower grades at each level. These variables are expected to have a negative impact on loans.
2. *The socioeconomic variables* used are a socioeconomic status index attached to the area in which the school is located and an urbanization index for the same

area. The socioeconomic status ranges from strongly underprivileged to privileged. We expect that the more privileged the group is, the higher the probability of loans and the number of loans per pupil are. An index of urbanization is also part of the model to characterize the students' environment. Although the influence of the level of urbanization on school library loans is expected to be significant, the sign of this influence is not clear a priori.

3. Students who have chosen to enter *professional programs* in the latter years of their secondary education are expected to use school libraries less than students in general programs, *ceteris paribus*.
4. *The proximity of a public municipal library* could have opposing effects on the use of school libraries by pupils. The first effect is one of substitution, students borrowing fewer books from school libraries because they use the public library. The second is that the habit of frequenting the public library could increase the students' demand for reading materials.

*The school characteristics* selected as relevant here can be divided into three categories:

1. *School size*, measured by the number of teachers, is expected to have a significant influence on lending policies.
2. *Whether or not a school is public or private, English or French*, or belongs to an *integrated school board* (regrouping both elementary and secondary schools), is expected to influence the lending policies of the school library; however, we do not have a priori anticipations on the direction of this influence.
3. *"Polyvalent"* secondary schools offer a wider range of options in their curriculum. This may influence the reading behavior of students and also the lending policies of the library.

One category of variables we had hoped to include in our causal model of library use was that of pedagogical variables. It seems clear that teachers have a strong influence on students' reading and therefore on school library loans. Information on teachers' pedagogical methods and on their ability to use the library materials efficiently is, unfortunately, not available. Nevertheless, as it stands, the model includes important explanatory factors and policy instruments and is particularly suitable for empirical analysis as will be seen in the following sections.

### *The Data and Construction of Variables*

The data used in this study came from two main sources: (1) a yearly school library survey yielding quantitative information on loans of printed materials, human resources, and physical facilities, and (2) a school data file providing information on the number of teachers and other school characteristics. For the academic year 1978-1979, this information was available for 877 public and private elementary school libraries and 350 public and private secondary school libraries, distributed across the province of Québec.<sup>5</sup>

To complement these data, the Canadian 1976 population Census and the Québec Municipality Library Repertory were used, from which were extracted area populations and information on the presence of municipal public libraries within reasonable distance of the schools. Furthermore, a socioeconomic index for the area in which each school in the sample is located<sup>6</sup> was constructed. Table 1 summarizes the information for the key variables of the model.

An interesting point is that most (80.8%) of the school libraries included in our sample make loans to students at the secondary level whereas less than half of them



(47.5%) do so at the elementary level. However for those libraries making loans, the number of annual loans per pupil decreases substantially from 38.0 at the elementary level to 17.0 at the secondary level. The average number of books and periodicals is larger at the secondary level but the variations in per pupil annual loans among school libraries are also higher at this level.

The quality of the stock of books, measured by the number of acquisitions and weedings in the previous year as a percentage of the total number of books, varies very little across levels of schooling or between subsamples; one exception is the somewhat surprisingly high rate (15%) of new acquisitions for the full sample of secondary school libraries. The figures on library personnel clearly show that this component is a very scarce resource indeed.

An average school library opens for substantially longer hours per week and provides more space and a larger number of carrels at the secondary than at the elementary level.

The three socioeconomic strata are fairly well represented in the sample but the four chosen levels of urbanization are not. School libraries in large urban areas seem under-represented. Finally, as expected, both larger schools (based on teacher number) and a higher fraction of private schools are found at the secondary level.

### *Empirical Results*

In discussing the results of the estimation of the model previously presented, a distinction is made between factors affecting the probability of a school library to make loans and the factors affecting the number of loans per pupil made over a school year. The statistical techniques used were probit analysis,<sup>7</sup> to estimate the probability of loans equation, and ordinary least squares, to estimate the number of loans equation. Results are discussed according to the same classification of variables as in the section describing the explanatory model.<sup>8</sup>

#### *The Quantity and Quality of Printed Materials*

The number of books per pupil is one of the most important determinants of the number of loans made by school libraries at both the primary and secondary levels. Furthermore, although there is some departure from uniformity in the results, the positive effect of the number of books is much stronger when library personnel is present. For instance, the interaction between the number of books and (i) the technicians working full-time at the elementary level or (ii) the professionals working part-time at the secondary level, increases the positive impact of the stock of printed materials on library use.

This indirect effect of personnel is confirmed by the results on the quality of books at the elementary level. Unless interacting with personnel variables, the quality of books (% of new acquisitions and culling) does not have a strong positive significant influence on the number of loans. A puzzling result is the lack of significant positive coefficients for quality and personnel interaction variables (and even one significant negative coefficient) at the secondary level.

As for the interaction between the number of books and physical facilities, it exerts a positive effect in the case of books and the library area and a negative effect in the case of books and number of chairs at the elementary level; at the secondary level, there is a positive effect in the case of books and number of carrels. It is not surprising that a larger area, facilitating a better presentation of books, should contribute to increase the circulation of additional books.

Turning to the determinants of the probability that a school library makes loans, we find that the number of books or periodicals has no direct (positive) effect in elementary or secondary schools. The number of books does exert a positive influence, however, through the interaction with personnel. The interaction between books and library area, at the elementary level, and between books and number of carrels at the secondary level also increases the probability that some loans are made. At the elementary level, one quality dimension, new acquisitions, has a positive direct effect which is reinforced by mostly positive effects for the interaction variables between acquisitions and personnel. At the secondary level, the other quality dimension, weeding, has a negative direct effect which is more than compensated for by a positive indirect effect (through the interaction variables with personnel).

The elasticities in Table 2 illustrate clearly the relative importance of the quantity and quality of printed materials in explaining (i) whether or not a library makes loans, and (ii) the number of loans per pupil made in a year. Elasticities give the change in percentage terms of the variables explained in the model following a percentage variation in any of the nondichotomized explanatory variables.<sup>9</sup> Use of the elasticities reveals that increasing the per capita number of books by 20% would lead to an increase of 10% in the number of loans per capita made by elementary and secondary school libraries. This is a strong reaction. The elasticity of the probability of loans with respect to weeding also appears to be quite high. However, as noted at the bottom of Table 2, approximate standard errors for elasticity estimates, in the case of the probability of loans, are large and therefore the results should be interpreted with some caution.

#### *Library Personnel*

An unexpected result of this study is the absence of any positive direct effect of personnel on loans at the primary level and the quasi-absence of such an effect at the secondary level. We have already discussed the generally positive indirect influence of personnel through the interaction with the size and quality of the stock of books. Such a positive indirect effect on the number of loans is also indicated by the positive coefficients of the interaction variable between the number of operating hours and personnel (part-time professionals and support staff) at the secondary level. On the other hand for other types of personnel (full-time professionals and part-time technicians), this interaction variable has a negative influence. Total elasticities of Table 2 summarize the information on the direct and indirect effects of library personnel on loans. The elasticities are rather small, which indicates that an increase in personnel would not lead to a substantial increase in loans, and that, moreover, some elasticities are negative.

#### *Library Physical Facilities and Organization*

The results on the effect on loans of the number of operating hours per week are as ambiguous as the results on personnel are inconclusive. The number of operating hours is a key element in explaining both the probability that a library makes loans and the number of loans made. The regression coefficients are positive and significant, especially at the elementary level, and the total elasticities of loans with respect to this variable are large. For instance, lengthening the hours of operation by 20% (going from 32.2 to 38.6 hours per week on average) would increase the number of loans by 17% in secondary schools.



Another organizational variable which has a significant influence is the presence of separate collections (outside the library). It reduced significantly the probability that an elementary school library makes loans. The physical facilities variables play a somewhat lesser role. The total library area per pupil increases the probability of loans at the elementary level and the number of loans at the secondary level. The number of chairs has a positive influence on library book loans while the number of carrels has the opposite influence at both schooling levels. A tentative explanation is that the presence of carrels may be an inducement to substitute in-library reading for borrowing.

### *User Characteristics*

The influence of socioeconomic status is significant but it is not uniform across schooling levels. Results show, at the elementary level, that the higher the socioeconomic index, the higher the probability of loans and the number of loans. On the contrary, at the secondary level, the higher the socioeconomic index, the lower the probability of loans and the number of loans.<sup>10</sup> The population size of the town or city where an elementary school is located influences positively the probability that the library makes loans but, for libraries already making loans, it reduces the number of loans per capita. The reverse is true at the secondary level although the pattern is somewhat less clear.

A higher proportion of students in the lower grades in the elementary and secondary schools increases the probability of loans. This unexpected result raises the question of the more or less intensive use of library materials by teachers at different grade levels.

The existence of a nearby municipal library influences negatively the probability of loans at both the elementary level and the number of loans at the secondary level. This might indicate that the actual relationship between public and school libraries is one of substitution rather than a complementary one for the younger clientele.

### *School Characteristics*

The larger size of the school, measured by the number of teachers, increases the probability of a library making loans but does not significantly influence the number of loans. Libraries of private secondary schools, most likely because of their specific vocational nature, make fewer loans. The fact that a secondary school is polyvalent (that is, it offers a wider range of options) has a negative effect on the number of loans.

### *Conclusion and Policy Recommendations*

This study illustrates the need to gather more and better data on variables associated with school library services. But even with incomplete information, the present study has shown that specific factors influence the use of the loan services offered by the school libraries.

The following conclusions and recommendations which policy makers should consider to complement their own personal judgements are worthy of note:

1. Among all variables, the number of books and the number of school library operating hours are the most important factors affecting positively the probability of loans and the number of loans made by school libraries.



2. Whether some types of library personnel are required to handle more books or have longer hours of operation is a technical question not addressed in the present study.<sup>11</sup> Although some positive indirect influence of personnel was indicated by the results, no direct influence of library personnel on the library output considered here was perceived.
3. For the physical facilities variables two results emerge: at the secondary level, the size of the library area has a positive influence on loans while, at the elementary level, the seating accommodation as represented by the number of chairs does so.
4. Some user characteristics and school characteristics influence the level of library services and this factor should be taken into account in designing policies.

In formulating policies to improve the economic efficiency of public and private school library policies, two further dimensions would have to be explored. First, one would want to know the influence of pedagogical methods and teachers' sensitization to the usefulness of school libraries on library use. Second, once a cost is assigned to each factor, a cost-benefit analysis should be performed so that public resources are allocated to library inputs efficiently. On that last matter, the nexus of number of schools, library personnel and costs must be seriously considered in light of the results obtained in this study.

#### Notes

1. Scattered references can be found in specialized library journals; for example, the *School Librarian* and the *School Media Quarterly*.
2. Some of those policy variables have been empirically considered in the public library literature; see, for example, Campbell and Shlechter (1979), Hayes (1979), and Powell (1978).
3. The reasoning is the same for "full-time professionals" whose influence can be reinforced by the number of "weeding." Formally, let  $Y$ : the number of loans per pupil,  $X_1$ : the number of books disposed of,  $X_2$ : the number of "full-time professionals." Assuming  $Y = b_1X_1 + b_2X_2$  and with  $b_1 = b'_1 + b''_1X_2$ , we obtain finally  $Y = b'_1X_1 + b''_1X_2X_1 + b_2X_2$ , where  $X_2X_1$  is the interaction variable. Now clearly this model cannot empirically be distinguished from the following:  $Y = b_1X_1 + b_2X_2$  with  $b_2 = b'_2 + b''_2X_1$  leading to  $Y = b_1X_1 + b'_2X_2 + b''_2X_2X_1$ .
4. This is also well documented in the public library literature; see, in particular, Stratton (1976) and D'Elia (1980a, 1980b).
5. Unfortunately, the city of Montreal, the largest urban community of the province of Québec, is under-represented in this sample.
6. Details and reference sources are given in Houle, Benyahia, and Montmarquette (1980), pp. 82-83.
7. Probit analysis deals with the problem of limited dependent variable as in the case when use of school library is defined equal to one and nonuse to zero in a regression framework; see Theil (1971, pp. 630-631) for details. D'Elia (1980b) has treated this problem in his analysis of use versus nonuse of public library services.
8. Complete regression results are available from the authors on request.
9. All the elasticities are calculated at the mean values of variables except for the elasticities of the probability of loans of secondary school libraries which are computed at the mean value of variables minus two standard errors (or set close to zero when a negative value results) to avoid being caught in a very flat portion of the probability curve, since most secondary school libraries do provide loans to their students.
10. Given the results, it is impossible to infer the existence of a substitution or complementary relationship between school library materials and reading materials available to the student at home.
11. On this particular problem, see Renner and Clark (1979).

This is a revised and condensed version of a report prepared for the Service général des moyens d'enseignement, ministère de l'Éducation du Québec, on *la mesure de l'impact des déterminants dans l'utilisation des services offerts par les bibliothèques scolaires*.

Financial assistance of the Centre de recherche en développement économique of the Université de Montréal is acknowledged for this revised version along with the technical assistance of Lina Di Pietro and Lisa Di Pietro.

References

- Campbell, D. E., & Shlechter, T. M. (1979). Library design influences on user behavior and satisfaction. *The Library Quarterly*, 47, 26-41.
- D'Elia, G. (1980a). A procedure for identification and surveying potential users of public libraries. *Library Research*, 2, 239-281.
- D'Elia, G. (1980b). The development and testing of a conceptual model of public library user behavior. *The Library Quarterly*, 50, 410-430.
- Hayes, R. M. (1979). The management of library resources: The balance between capital and staff in providing services. *Library Research*, 1, 119-142.
- Houle, R. (1980). *Rapport d'enquête sur les habitudes de lecture des élèves du secondaire*. Québec: Collection SREP, Gouvernement du Québec.
- Houle, R., Benyahia, H., & Montmarquette, C. (1980). *La mesure de l'impact des déterminants dans l'utilisation des services offerts par les bibliothèques scolaires*. Montréal: Centre de recherche en développement économique, Université de Montréal.
- Powell, R. R. (1978). An investigation of the relationship between quantifiable reference service variables and reference performance in public libraries. *Public Libraries*, 48, 1-19.
- Renner, C., & Clark, B. M. (1979). Professional and nonprofessional staffing patterns in departmental libraries. *Library Research*, 1, 153-170.
- Stratton, P. (1976). *Public libraries: Their structure, use and cost of service provision*. De Kalb, IL: Northern Illinois University.
- Theil, H. (1971). *Principles of econometrics*. New York: John Wiley & Sons.

NORMAN ROBINSON

*Simon Fraser University*

and

CAROLINE STACEY

*Delta School District*

## Serving on the School Board: A Political Apprenticeship for Higher Office

*The following question was investigated in this study: How do successful politicians who began their political careers as school board members and who have gone on to higher political office evaluate their school board experience in terms of its value as a political apprenticeship? Data were gathered from structured interviews with ten federal and provincial politicians who had once served as school board members; they agreed unanimously that school board experience was a valuable political apprenticeship for higher office. This was true regardless of the individual's political socialization or political recruitment experiences prior to school board service. The politicians felt that they had acquired valuable political skills and that their school board experiences had a strong politicizing effect on them. During their period of service on the board they had experiences which changed their political ambitions from office-bound ambitions (at the school board level) to upward-bound ambitions (at the federal or provincial level).*

Researchers have given little attention to investigating the value of the school board experience as a political apprenticeship for those individuals who leave the school board and go on to higher political office. The reason for this neglect probably lies in the fact that the great majority of school board members do not have political ambitions above the school board level. For example, Ziegler, Jennings, and Peak (1974) found that only 7% of incumbent school board members had upward-bound political aspirations. In an earlier study which investigated the rea-

---

Norman Robinson is a professor of educational administration at Simon Fraser University. His major research interest is the politics of educational governance.

Caroline Stacey is a recent graduate of the Administrative Leadership Program at Simon Fraser University. She is currently working as a school administrator in Delta School District in British Columbia.



sons why people run for school board, McCarty (1959) found that less than 10% of school board candidates listed future political advancement as a motive for running. In a replication of the McCarty study done in British Columbia, McDonald (1972) found a similarly small percentage of candidates running to further their political careers.

The question of how many school board members actually do go on to higher political office has not been thoroughly researched. Existing evidence suggests that probably more school board members in Canada move on to higher political office than is the case in the United States. For example, Schlesinger (1966) found that few candidates for governorship and the U.S. Senate had a background of local government experience. In Canada, on the other hand, studies (Kornberg, 1967; Kornberg & Mishler, 1976; Kornberg, Smith, & Clarke, 1979) done on the backgrounds of members of the House of Commons reveal considerable experience in local government.

School board experience, as distinct from other forms of local government experience, has not been isolated in either the American or Canadian studies. As a result, no complete and accurate information exists on how many former school board members have gone on to higher political office in the United States or Canada. An inspection of the most recent Canadian parliamentary guide (Normandin, 1982) does reveal, however, that 23 of the present 282 members of the House of Commons are former school board members. About ten percent of the members of the provincial legislatures are former board members. Ontario, for example, has 17 former school trustees in its 125 member house.

Little also is known about school board members who seek higher political office but are unsuccessful in their efforts. No complete and accurate information on this question has yet been assembled. A recent study by Stacey (1983) did, however, document the number of school board members in British Columbia who offered themselves for election in the Canadian federal election in 1979 and in the provincial elections in 1979 and 1983. In the federal election, a total of six former school board members sought election to B.C.'s 28 seats in the House of Commons. Three of these former school board members were successful and became members of parliament. In the provincial election of 1979, there were 16 former school trustees seeking election to the 57 seats in the B.C. legislature; four were successful in winning seats. In the 1983 B.C. provincial election, 13 former trustees sought election and four were successful.

What becomes clear is that a small but important number of former school board members are successful in their quest for higher political office, some of whom later rise to fill high level political offices.

### *The Research Question*

The research question that this study attempted to investigate can be stated as follows: How do successful politicians who began their political careers as school board members, and who have gone on to higher political office at the federal or provincial level, evaluate their school board experience in terms of its value as a political apprenticeship?

In studying this question, an attempt was also made to examine the political antecedents of the school board experiences of the politicians studied. That is, an examination was made of the political socialization experiences and the political recruitment experiences leading up to the politicians' school board experiences.

## *Methods*

### *The Sample*

The population for this study consisted of all former school board members, living in B.C., who are serving or have served in the House of Commons or the Legislative Assembly of British Columbia. In spite of a rather thorough search of the Canadian parliamentary guides over the past five years, it was not possible to identify the exact size of this population and to come up with an appropriate sampling procedure. As a result, a decision was made to interview all of those who could be contacted and who were willing to be interviewed. This turned out to be ten people.

In research terms, this sample of ten politicians is very much an accidental sample—it is not a random sample. It does possess, however, great diversity in terms of political characteristics which reflects the reality of political life in B.C. The sample had the following characteristics:

1. it contained representatives from all four political parties currently active in B.C.—four New Democrats, two Liberals, two Progressive Conservatives, and two Social Crediters;
2. five of the politicians had service in the House of Commons—four in the B.C. Legislature and one in both;
3. three of the politicians had service in either a federal or provincial cabinet;
4. three of the politicians had retired from politics; seven were still active;
5. four of the politicians had served only on the government side, three only on the opposition side, and three on both sides of the house;
6. there were eight men and two women; and
7. the period of service in political office at the federal or provincial level ranged from a low of two years to a high of 26 years.

### *Interviews*

Structured in-depth interviews were used in this study. The interview schedule included questions regarding the political socialization, political recruitment and political apprenticeship dimensions of this study. The principal questions were as follows:

1. What were the important factors in the early *political socialization* of these politicians?
2. What were the important factors in the *political recruitment* process that led up to their school board service?
3. In what ways can service on a school board be considered a *political apprenticeship* for higher public office?

## *Results*

### *Early Political Socialization*

In political science research, political socialization is usually viewed as the process by which an individual acquires political values, attitudes, and interests. In addition, the individual acquires a knowledge of the political community with its institutions and actors. From childhood onward political socialization is a continuous process with the individual being placed under constant potential change through the influence of family, peers, school and other sources.

It was important in this study, first, to determine what factors played a significant part in the early political socialization of the ten politicians studied and, secondly, to assess the effect that these socialization experiences had on the initiation and development of the political careers of these ten individuals.

### *Findings*

Family experiences played an important part in the early political socialization of eight of the ten politicians. Only two of the politicians came from families where there was no active interest in politics. For the other eight, an interest in politics was characteristic of family life. Typical of the comments describing the family situation were the following:

My parents were both very active politically. They talked politics at the dinner hour . . . so I grew up with it.

My father and mother were both actively involved in politics. My mother mainly in a supportive role, my father mainly in the union, as president of his local. He was deeply involved in union politics, and as I said, ran for alderman in the city.

Four of the politicians came from family situations where a parent or close relative had held public office.

I came from a very political family. My father was a Tory and he served on the school board too. The old time Tories used to meet at our house with my father and I used to love to hear them talk.

My father was on the school board for several years. He was chairman. Later he was elected to the House of Commons.

The interviews with the politicians revealed as well that the majority of them (six of the ten) had become politically active in a formal party well before their entry into school board affairs—a few in adolescence and others, after high school, usually in university. Typical of the experiences described are these three comments:

My first public experience was when I was in grade eight and I was asked to give the toast to the graduating class. From then on I became active in student council affairs in high school. I joined the C.C.F. at a very young age . . . so at the age of fourteen or fifteen I was attending political meetings.

I had been involved in the Liberal Party since university days and followed through afterwards . . . I remained an active member and was president of a very small local at the time I became a school trustee.

My father ran for alderman a number of times . . . and I delivered campaign literature for him. I joined the party after that. So I've always been involved in politics in some way, at university and elsewhere. At the time I ran for trustee I was president of the local N.D.P. club.

### *Discussion*

The findings of this study reveal quite clearly that the majority of the politicians interviewed came from politically active families. In most cases, politics was part of family life and in nearly half of the families, either a parent or close relative had held public office.

These findings support those of other studies which show that the level of political interest and activity is much higher in the families of future politicians than is the case among members of the general public. A study by Kornberg and Mishler (1976) of Canada's 28th Parliament (1972-1974) showed that nearly 66% of the



members had fathers with a strong interest in politics and 38% had mothers with a similar high level of interest. Furthermore, Clark, Price, and Krause (1976), in a study of provincial legislators in all ten Canadian provinces, found that 35% had parents or other relatives who had held public elective office.

What the findings of this study do not support is the conclusion of Zeigler, Jennings, and Peak (1974) that school board members do not come from politicized family backgrounds. Their study of 490 school board members in 83 American school districts showed that only one-third of the board members came from families where the father was "very much interested" in politics and only one-fifth of board members put their mothers in that category.

What is quite clear from the findings of this study is that the overwhelming majority of these future school board members underwent early political socialization experiences that were strong and pervasive. What is not clear is whether most individuals who seek school board membership have similar kinds of political socialization experiences. Are strong and pervasive early political socialization experiences found only in the backgrounds of board members whose political ambitions become upward-bound (i.e., oriented to attaining higher political office) or are they found as well in the backgrounds of board members whose political ambitions remain office-bound (i.e., restricted to the school board)? As well, it is unclear whether there are differences between Canada and the United States in terms of the political socialization experiences that prospective office seekers undergo prior to standing for election to a school board.

### *Political Recruitment to the School Board*

The fact that an individual comes from a highly politicized family environment and has strong and pervasive political experiences in youth and early adulthood does not mean that an individual will embark on a political career. In any society there are thousands of individuals who come from highly politicized backgrounds but only a few ever stand for public office. Thus, Milbrath (1965) and Bowman and Boynton (1966) suggest that future political actors have to cross a political activity threshold. They have to be stimulated to stand for public office either through the encouragement of others or through the arousal of their own inner feelings. This crossing of the political activity threshold constitutes the process of political recruitment. In this study, the focus was on what factors were important in the political recruitment process that lead up to service on the school board by these politicians.

### *Findings*

Election to the school board constituted the beginning of a political career for nine of the ten politicians. Only one had prior experience in a public elective office.

Prior to their election to school board, three of the politicians were nonactive in community affairs; the remaining seven had a varied pattern of community activity. In their study of school board members, Zeigler, Jennings, and Peak (1974) examined the community activities of members prior to their election to office. These researchers classified the community activities into three categories: (1) civic/business activities (i.e., involvement in service clubs or business groups); (2) political/governmental activities (i.e., political party involvement or service on public boards); and (3) educational activities (i.e., involvement in educational affairs as a teacher, administrator, or member of a parent-teacher group). Quite clearly these are not dis-

crete categories and it is not possible to categorize the community activities of the individuals in this study into three tidy areas. It was possible, however, to identify from the interviews what the primary focus of their community activities was. For the seven who were active in community activities, two can be classified as active in civic/business affairs and five in political/governmental affairs. None could be classified as being active in education in spite of the fact that three were former school teachers. In all three of these cases, the individuals involved had a long history of political involvement and party membership before they entered teaching.

A number of writers (Zeigler, Jennings, & Peak, 1974; Goldhammer, 1964) have pointed out that a disproportionately high number of school board members come from "education" families, that is, families where parents or relatives are employed in educational positions. This is certainly true of this study. As has been noted already, three members had themselves been former teachers, one had been an educational administrator, and four others had close relatives in educational positions. This meant that only two of the ten politicians had no personal or family involvement in education.

In terms of their recruitment to stand for election to school board, the politicians interviewed were asked to name persons or groups who encouraged their candidacies. Six had been approached and encouraged to run, the other four indicated that their candidacies were self-initiated. Of the six who had been approached, three were approached by informal groups (parents, friends, neighbors), two were approached by political leaders (mayors in both cases), and one was approached by a formal group (ratepayers' association).

Issue differences with the incumbent school board played an important role in stimulating the school board candidacies of six of the ten politicians interviewed. One whose candidacy was self-initiated said:

I was always very vocal, it was part of my background. And one day I was sitting in the staff room complaining about the school board and somebody said, "You should get on that school board and then maybe you would stop complaining so much about it." So after I left teaching I decided to run for trustee. I became a trustee because there were things going on in education in our district that I didn't like.

Four of the ten politicians who decided to run for school board did not have issue differences with the incumbent board members. Their reasons were often highly personal. One said:

After the (federal) election, six of us sat around and conducted a post mortem on why we hadn't been able to get anyone elected. I suggested that we should start seeking municipal office . . . so that we would have a track record that would indicate that we could hack public life. . . . So I decided to run for school board and eventually I became an M.P. . . . I planned a career in politics, beginning with the school trusteeship.

Kornberg and Mishler (1976) did an analysis of the motives of candidates who sought office and were successful in becoming members of parliament in the Canadian federal election of 1972. Although there was much overlap in categories, they found that the candidates were motivated by the following (in order of importance): (1) "good government" considerations; (2) ideological considerations; (3) "challenge of office" considerations; (4) party pressure; and (5) pressure from others.

In this study the six candidates who were approached to run for school board tended to have different motives for running than the four individuals whose candidacies were self-initiated. The motives of those approached could be roughly categorized as follows: (1) pressure from others (three candidates); (2) "good govern-



ment” considerations (two candidates) and (3) ideological considerations (one candidate). For the four self-initiated candidates, the breakdown on motives is as follows: (1) “challenge of office” considerations (three candidates); and (2) ideological considerations (one candidate).

### *Discussion*

The pattern of political recruitment to school board service for the individuals in this study differs considerably from the pattern of recruitment for school board members reported elsewhere (Goldhammer, 1964; Jennings & Zeigler, 1971; Cistone, 1974; Zeigler, Jennings, & Peak, 1974). The two most extensive studies of the political recruitment of school board members are those of Cistone (1974), done in Ontario, and Zeigler, Jennings, and Peak (1974), done in the U.S.A. on a nationwide basis.

An exact comparison cannot be made between these three studies because each study did not ask the same questions in all areas. The limited comparisons that can be done, however, do reveal certain important differences between the results of this study and the results of the earlier studies of Cistone and Zeigler, Jennings, and Peak. In terms of political recruitment to the school board, the results of this study show a higher level of political/governmental activity in the community prior to school board service, more self-initiated candidacies, no encouragement to run from incumbent school board members or professional educators, and more issue differences with incumbent board members.

The only major similarity in the results of this study and the results of the Cistone and the Zeigler, Jennings, and Peak studies is that all three show that school board members are drawn predominantly from education families. Over 80% of board members in all three studies had relatives in education and a sizeable minority of board members had been professional educators.

One finding of this study that deserves particular comment relates to the total absence of any activity on the part of incumbent school board members to approach or encourage potential candidates to run for office. A number of writers (Crain, 1968; Goldhammer, 1964; Jennings & Zeigler, 1971; Tucker & Zeigler, 1980) have noted that there is a strong tendency on the part of school boards to perpetuate themselves. This is done by encouraging the candidacies of carefully selected individuals. In this study, none of the politicians interviewed indicated that they received any support or encouragement from incumbent school board members to run for office.

### *School Board Service as a Political Apprenticeship*

In this study an attempt was made to determine whether school board experience served as a political apprenticeship for the successful federal and provincial politicians studied, and if so, in what ways the experience was valuable to them.

### *Findings*

All ten politicians were of the strong opinion that their school board experience was an invaluable political apprenticeship. For all but one of the politicians, service on the school board was their first experience in public office and for them it was a powerful and mind-expanding experience. As one of the politicians described it:

I look back on my school board years as one of the most interesting and stimulating times of my life. It was my first entrance into politics and everything was to my mind so excit-



ing. I was almost overwhelmed with the fact that I had been given the opportunity to sit on the board and to make decisions.

The politicians mentioned particular skills they acquired as a result of their school board experience. Foremost among these were skills in dealing with people, particularly in the areas of coalition building and conflict resolution. One of the politicians put it this way:

You may have the best ideas in the world as a school trustee but until you get half of the people on your side, it won't go anywhere. So you have to work with people whom you may not even care to talk with, with whom you have nothing in common, but with whom you have to work with in order to accomplish what you want. As a politician, learning that when I was a trustee was very useful.

In describing the human relational skills they acquired as a result of their school board experience, other politicians said:

You learn to work in groups, how to come to decision making, how to share strong ideas . . . I think it's excellent political experience.

You become adept at . . . juxtaposing conflicting roles and dealing with people who may or may not have some opposing view.

Several of the politicians interviewed made specific reference to the skills they acquired in dealing with professional experts. One politician said:

I learned not to have blind faith in expertise. A lot of teachers, administrators and experts didn't know what they were talking about. I learned a built-in suspicion to the expert as opposed to the common sense approach of people.

Another politician added,

I learned to be skeptical with experts and with conclusive findings.

All of the politicians expressed the view that they acquired valuable skills in financial matters. Typical of the comments relating to the acquisition of financial skills is the following:

I learned to be resourceful. How to husband monies and how to zero in on the source of a problem and how to then go about solving it.

The political skills of coalition building, conflict resolution, assessment of expert evidence, and financial management were identified by the politicians as the most valuable skills they acquired as a result of their school board experience. Certain other important political skills that were also acquired but which were mentioned less frequently include skills in chairing meetings, communicating to groups, listening to others.

Another important finding of this study was that school board experience had a politicizing effect on the individuals involved. They found that they liked politics, they could be successful at it, and this interest and success encouraged them to consider running for higher political office. One of the politicians put it succinctly by saying:

People either become political animals or they don't after serving on a board . . . I became one in the sense that I enjoyed politics.

Another politician described how success in the school board experience can serve as a stimulus for the seeking of higher political office:

If you are successful as a trustee you are motivated to go on. . . . If you are successful in improving society in whatever area, then it's only natural that if you see you can contribute in some larger way you will be motivated to continue.

With the exception of one individual, none of the politicians consciously planned to use the school board as a stepping stone to higher political office. The stirring of political ambition came about as a result of the school board experience. One politician described the process as follows:

Most of the people I know that have come from school boards and (city) councils didn't enter the council or trustee job thinking they would run for higher office. I think they sought the job and that was the end to the goal. As they got into it, they enjoyed it and perhaps became popular in their own area, be it in their community or their political party, and saw the opportunity present itself one day to go on. I don't think that people who go out for school board and council really see it as a means to an end. I think they go out to change the world, and are either in the end defeated, drop out, or go on. It's like any other process, there is room to keep on climbing. Any person who is ambitious, when they start on a job, will be aiming for a better job.

Success and a sense of accomplishment in the school board member role was thus of great importance in encouraging these individuals to seek higher public office. Three of the individuals interviewed did not feel, however, that they had accomplished as much as they had wanted. They felt that this was less a function of themselves and more a function of the limited power school boards have. One of the former school board members put it this way:

My feeling was that when I went in and became a school trustee most of the decisions had already been made by the time they were presented to us at the school board level. The government makes decisions on financing and there is very limited room to move as a trustee. I think somebody once said that fifteen percent of the budget is yours to do whatever you want with, but only if you raise taxes. I don't think that it is a particularly satisfying job.

An examination was also made of the effect that electoral defeat had on the political ambitions of these former school board members. Only two of the ten politicians had ever suffered electoral defeat, and only one had been defeated in a school board election. The effects of defeat tend to be temporary and no real antidote to political ambition. The two politicians in this study who had experienced defeat described their reaction as follows:

Well, I tangled with some teachers and was defeated the second term. That was it, I had had enough. Later, however, I was asked to run for city council and eventually became mayor.

Yes, I lost two elections. I learned a lot by losing. When you understand why you lost, you can do something about it. . . . It's depressing, but not devastating, to lose. You have to understand that most people run to win. You are perhaps depressed for a week or so, but my God, it's only an election!

Finally, an attempt was made to determine whether school board experience made the individuals in this study desirable to political parties as potential federal or provincial candidates. This was a difficult question to sort out as many of the individuals were engaged simultaneously in school board affairs and party affairs at the constituency level. Several of the politicians did indicate, however, that their school board experience was very important in terms of their desirability as a candidate for higher office. One said:

If I hadn't been an elected official for ten years I'm sure that members of my party would not have approved me to run. A party wants to win and when they are looking for someone, they are looking for someone who is electable.

Another described his experience in these terms:

I think the fact that I was a school trustee influenced people to ask me to seek further office. I had always been involved in politics, long before I was a school trustee but basic-



ally at the party organization level rather than in elected office. Getting elected as a school trustee introduced me to a whole new realm of politics. Then when the former candidate decided not to run again, Tommy Douglas approached me, and the local party president approached me, and I said "sure I'll make a go for it."

### *Discussion*

To this time, researchers have given little attention to school board service as a political experience. What the results of this preliminary study show is that it can be a powerful experience in terms of the acquisition of political attitudes and skills. Regardless of the nature of their earlier political socialization or political recruitment experiences, the individuals in this study were unanimous in their agreement on the value of the school board experience as a political apprenticeship. They became more politicized as a result of the experience, and in addition, acquired skills which they deemed to be invaluable in their future political careers.

### *Summary and Implications*

This study examined the value of the school board experience as a political apprenticeship for a group of ten politicians who went on to successful careers at the federal or provincial level. These politicians agreed unanimously that school board service was a valuable political experience for them. It was one in which they acquired needed political skills and attitudes. They also acquired a certain desirability as potential candidates for higher public office.

Only one of the ten politicians in this study deliberately set out to build a political career using the school board as a stepping stone. The remainder did not have upward-bound political ambitions at the time they became board members—they simply wanted to be school trustees. Yet during the course of their tenure as board members they had experiences which changed their political ambitions from office-bound ambitions to upward-bound ambitions. This was true regardless of the nature of their prior political socialization or political recruitment experiences.

In this study we examined the school board experiences of individuals who have gone on to higher political office. We also examined their political socialization and political recruitment backgrounds. Most school board members do not go on to higher office. Quite a few try, but most fail. For everyone who serves on a school board, the experience is, among other things, a political experience. For all the individuals in this study, the experience was useful and for some it was also powerful and mind-expanding. For those board members who do not go on to higher office, or for those who fail in their efforts to do so, the nature of school board service as a political experience may well be different. It may be that their political socialization and political recruitment experiences differ markedly from those of the politicians in this study. Only further research can clarify these issues and our future research efforts will focus on these issues.

### *References*

- Bowman, L., & Boynton, G. R. (1966). Recruitment patterns among local party officials. *American Political Science Review*, 60, 667-676.
- Cistone, P. J. (1974, October). School board member recruitment: The case of Ontario. *Journal of Educational Administration*, 12(2), 43-56.



- Clarke, H. D., Price, R. G., & Krause, R. (1976). *The role socialization of the Canadian political elite: A note on provincial legislators*. Unpublished paper, University of Windsor.
- Crain, R. L. (1968). *The politics of school desegregation*. Chicago: Aldine.
- Goldhammer, K. (1964). *The school board*. New York: Center for Applied Research in Education.
- Jennings, M. K., & Zeigler, H. D. (1971). Response styles and politics: The case of the school boards. *Midwest Journal of Political Science*, 15, 290-321.
- Kornberg, A. (1967). *Canadian legislative behavior*. New York: Holt, Rinehart & Winston.
- Kornberg, A., & Mishler, W. (1976). *Influence in Parliament: Canada*. Durham, NC: Duke University Press.
- Kornberg, A., Smith, J., & Clarke, H. D. (1979). *Citizen politicians: Canadian party officials in a democratic society*. Durham, NC: Carolina Academic Press.
- McCarty, D. J. (1959, September). School board membership: Why do citizens serve? *Administrator's Notebook*, 8(1).
- McDonald, K. R. (1972). *The interrelationships of motives for seeking school board membership, mode of school board operation and community type*. Unpublished master's project, Simon Fraser University.
- Milbrath, L. (1965). *Political participation*. Chicago: Rand McNally.
- Normandin, A. L. (Ed.). (1982). *Canadian Parliamentary guide*. Ottawa: Editing and Administration, P.O. Box 3453, Station C.
- Schlesinger, J. (1966). *Ambition and politics: Political careers in the United States*. Chicago: Rand McNally.
- Stacey, C. (1983). *The school trusteeship as a political experience*. Unpublished master's project, Simon Fraser University.
- Tucker, H. J., & Zeigler, H. D. (1980). *The politics of educational governance: An overview*. Eugene, OR: ERIC Clearinghouse on Educational Management.
- Zeigler, L. H., Jennings, M. K., & Peak, G. W. (1974). *Governing American schools*. North Scituate, MA: Duxbury Press.

FRANK M. MERRITT

and

R. STEVE McCALLUM

*University of Southern Mississippi*

## The Relationship Between Simultaneous-Successive Processing and Academic Achievement

*The present study investigated the relationship between simultaneous-successive information processing and academic achievement among 157 college students. A factor analysis of 6 tasks comprising the simultaneous-successive battery yielded a two-factor solution. Four processing groups were formed using factor scores as criteria. Cumulative grade point average (GPA) and composite American College Testing (ACT) scores represented the dependent (achievement) variables. Results indicate that level of simultaneous and successive processing is related systematically to GPA; high levels of both processing modes are necessary for high GPA achievement. On the other hand, simultaneous processing seems relatively more important for high ACT performance. These results are discussed as they relate to task demands of the ACT and GPA.*

The study of human learning often progresses by using scientific models consistent with the existing Zeitgeist. For example, several current cognitive models utilize computer programming to simulate learning. Similarly, increasingly sophisticated neurophysiological research has made neuropsychological models more relevant. Recently, these two types of models have been integrated, creating a neuropsychological model of human learning which is typically described in computer processing terminology. Founded upon A. R. Luria's clinical observations (Luria, 1966a, 1966b), this model has been elaborated and extended by J. P. Das

---

Frank M. Merritt is currently at the University of Southern Mississippi Psychology Department completing doctoral work in Clinical Psychology. He maintains active research interests in neuropsychologically-based learning models.

R. Steve McCallum is currently an Associate Professor of Psychology at the University of Southern Mississippi and Associate Editor of the *Journal of Psychoeducational Assessment*. He also maintains active research interests in neuropsychologically-based learning models.

and colleagues via empirical research (see Das, Kirby, & Jarman, 1979). This model is often termed the Luria-Das Information Processing Model.

The structural and processing features of the Luria-Das model can be briefly described as follows. The brain is composed of three units. Unit 1 is primarily composed of upper/lower brain stem organs and is responsible for the arousal or motivational function; Unit 2 is located in the posterior region of the neocortex including the occipital, temporal, and parietal lobes. Unit 2 structures receive information, analyze it into elementary components, and code it into a symbolic/physiological form useful to the entire brain. Information is analyzed and coded within Unit 2 in either a simultaneous or a successive fashion. Simultaneous integration refers to the synthesis of separate elements into groups, often with spatial characteristics. Any part of the gestalt and its relative position is immediately surveyable. On the other hand, successive integration refers to the processing of information in a serial order so that information is surveyable only in a temporal, orderly manner, with knowledge of each bit of information dependent upon the previous bit. Finally, Unit 3 is comprised of the prefrontal lobes, and is responsible for planning, executing, and evaluating problem-solving strategies.

Conceptualizing human learning within this comprehensive model should provide guidance in the selection, validation, and interpretation of measurement instruments. Contrasted to traditional empirical approaches to intellectual assessment which are primarily based on predictive validity, the Luria-Das model utilizes relevant theoretical and empirical research to conceptualize cognitive functioning. Therefore, results from testing are amenable to interpretation within a meaningful conceptual framework and subsequent remedial strategies are easily established (e.g., Kaufman & Kaufman, 1979; Krywaniuk & Das, 1976).

Obviously, empirically derived tests which predict achievement are valuable to educators. Tests yielding good predictive validity *and* information relevant for interpreting performance would be even more valuable. Studies exploring the relationship between simultaneous-successive processing described within the Luria-Das model and academic achievement offer evidence that such tests may be possible in the near future. Results from elementary school age children indicate that simultaneous-successive processing and achievement are related, but in a complex way. For example, when a simultaneous-successive battery is administered along with reading achievement tests, results reveal that successive processing is the mode often employed by beginning readers; more sophisticated readers seem to rely on both modes (Das, Kirby, & Jarman, 1979).

The present study was designed to investigate the relationship between simultaneous-successive processing and academic achievement in college students. Questions focused on whether simultaneous-successive processing ability varied systematically with achievement. Individuals demonstrating high simultaneous and successive processing skills were expected to earn the highest achievement scores whereas those demonstrating low simultaneous and low successive processing skills were expected to earn the lowest achievement scores; consistent with the non-hierarchical nature of the model (Das et al., 1979; Jarman, 1978), individuals demonstrating high simultaneous-low successive ability should perform similarly to those exhibiting low simultaneous-high successive ability.



## Method

### Subjects

Participants included 157 juniors and seniors enrolled in Educational Psychology courses at a mid-sized southeastern university. This sample was composed of 53 males, 104 females; 133 were white, 24 black. Ninety-seven were education majors. Data from males and females were combined for analyses (see Merritt & McCallum, 1983; Kirby, cited in Das et al., 1979, for justification).

### Simultaneous and Successive Tasks

Six tasks commonly used to assess simultaneous-successive processing were administered in group format. The tasks included:

*Raven Progressive Matrices.* This test is described as a culture-reduced measure of nonverbal reasoning (Raven, 1958). Participants indicated which of a given set of alternatives correctly completes a visual pattern. Fifteen of the 60 available items (every fourth item) were chosen for presentation.

*Memory for Designs.* The nine items from the Bender Visual Motor Gestalt Test (Bender, 1946) were presented for four seconds each. Following the presentation, examinees constructed the figures from memory.

*Figure Copying.* The last 12 figures from the Developmental Test of Visual Motor Integration (Beery & Buktenica, 1967) were presented for 30 seconds each and copied during that time.

*Free Recall.* Ten lists containing six monosyllabic words were presented orally. All 60 words appeared at the 1st, 2nd, or 3rd grade levels of the Slosson Oral Reading Test (Slosson, 1963) or the Basic Sight-Word Test (Dolch, 1942). The order of recall was not critical.

*Serial Recall.* Ten additional lists of 6 monosyllabic words from the 1st, 2nd, and 3rd grade levels of the Slosson Oral Reading Test and the Basic Sight-Word Test were presented orally. The order of recall was critical.

*Digit Span-Forward.* This task required recall of increasingly longer lists of digits. The 14 lists ranged from 4 to 10 digits in length. Each was recalled and written immediately after oral presentation.

### Rationale for Task Selection

Of the six tasks, Raven Matrices, Figure Copying, and Memory for Designs are considered indices of simultaneous processing. Serial Recall, Free Recall, and Digit Span-Forward are purported to be measures of successive processing. Evidence for construct validity of the battery comes from several sources. Luria's clinical observations of brain-damaged individuals clearly indicate the existence of these two modes of processing, and the six tasks described above exhibit face validity; i.e., they appear to assess the two basic processing modes described by Luria. And though the simultaneous-successive tasks have been psychometrically classified into parallel categories such as nonverbal-verbal, visual-auditory, and reasoning-memory, the two factors which have emerged from analyses can be appropriately termed simultaneous and successive processing. For example, simultaneous processing is utilized in verbal tasks such as the perception of lexical ambiguity (Das & Jarman, 1980), in auditory tasks such as syllogisms (Cummins, cited in Das, Kirby, & Jarman, 1975), and in memory tasks such as memory for designs (McCallum & Merritt, 1983). Additional convergent and discriminant validity is provided by Das et al. (1979).

### *Achievement Measures*

Achievement measures were cumulative grade point average (GPA) and the American College Testing (ACT) composite score. The ACT composite score reflects assessment of English and Math usage, and Social Studies and Natural Science Reading skills; it is typically administered to high school seniors and valued for predictive validity of college success. The scores were obtained from student records kept by the university.

### *Procedure*

Before the simultaneous-successive battery was administered, a questionnaire requesting demographic data was completed by each student. The six tasks were then administered in group form, with simultaneous tasks alternating with successive tasks. The order of presentation of the tasks shifted systematically with each presentation. Administration was standardized using a slide projector and taped audio, and required approximately 55 minutes.

Task scores were subjected to a principal components analysis; factors with eigenvalues greater than 1.0 were rotated according to a varimax criterion. Grouping of subjects was accomplished on the basis of mean splits of factor scores. Therefore, group assignments reflect relative ability to process information. Performance of the group members were compared on ACT scores and cumulative GPA, the dependent variables. Contrasts predicted from the model employed one-tailed tests of significance.

### *Results*

As in a previous study employing a college student sample (McCallum & Merritt, 1983), the factor analysis of task scores yielded two factors labelled simultaneous processing (defined by Raven Matrices, Memory for Designs, and Figure Copying) and successive processing (defined by Free Recall, Serial Recall, and Digit Span-Forward). (See Table 1 for factor loadings, means, and standard deviations.) Factor scores were subsequently obtained and became the criteria for assignment to one of four groups—high simultaneous-high successive (HiSim-HiSuc), high simultaneous-low successive (HiSim-LoSuc), low simultaneous-high successive (LoSim-HiSuc), and low simultaneous-low successive (LoSim-LoSuc). Because group assignments were based on mean splits of factor scores, groups reflect *relative* ability to process information.

As previously mentioned, in order for performance to be consistent with the model, members of the HiSim-HiSuc group should have earned significantly higher achievement scores than members of any other group. Conversely, members of the LoSim-LoSuc group should have earned scores significantly lower than those from any other group. These ten contrasts were predicted by the model and allowed use of a priori *t* tests.<sup>1</sup> Performance of the group members within the two intermediate or “off-diagonal” groups should have been similar because the model suggests a nonhierarchical arrangement of the two processing modes. That is, the two groups high on only one processing dimension should have performed about equally. These two contrasts employed two-tailed criteria. Means, standard deviations, and *t* ratios for all group comparisons are presented in Table 2.

Results employing the cumulative GPA were consistent with predictions from the model for all contrasts. Table 2 presents the relevant values for the cumulative GPA analyses.

TABLE 1  
MEANS, STANDARD DEVIATIONS, AND FACTOR LOADINGS  
FROM SIX COGNITIVE TASKS

| Tasks              | <u>M</u> | <u>SD</u> | Successive | Simultaneous |
|--------------------|----------|-----------|------------|--------------|
| Digit Span         | 8.18     | 2.21      | .85        | -.02         |
| Free Recall        | 49.87    | 5.54      | .76        | .34          |
| Serial Recall      | 43.57    | 7.43      | .81        | .24          |
| Raven Matrices     | 8.08     | 2.10      | .11        | .77          |
| Memory for Designs | 28.14    | 1.51      | .22        | .58          |
| Figure Copying     | 7.24     | 2.24      | .10        | .85          |

TABLE 2  
COMPARISON OF THE MEANS OF CUMULATIVE GPA AND COMPOSITE ACT  
FOR INFORMATION PROCESSING GROUPS

| GROUPS      | ACT      |           |          |           | GPA      |           |          |           |
|-------------|----------|-----------|----------|-----------|----------|-----------|----------|-----------|
|             | <u>M</u> | <u>SD</u> | <u>t</u> | <u>df</u> | <u>M</u> | <u>SD</u> | <u>t</u> | <u>df</u> |
| HiSim-HiSuc | 18.92    | 3.80      | .07      | 63        | 3.04     | .48       | 2.47**   | 82        |
| HiSim-LoSuc | 18.85    | 4.58      |          |           | 2.77     | .50       |          |           |
| HiSim-HiSuc | 18.92    | 3.80      | 3.35**   | 60        | 3.04     | .48       | 2.21*    | 77        |
| LoSim-HiSuc | 15.70    | 3.43      |          |           | 2.79     | .49       |          |           |
| HiSim-HiSuc | 18.92    | 3.80      | 4.73**   | 63        | 3.04     | .48       | 4.37**   | 90        |
| LoSim-LoSuc | 14.38    | 3.77      |          |           | 2.56     | .56       |          |           |
| HiSim-LoSuc | 18.85    | 4.58      | 2.69**   | 47        | 2.77     | .50       | .15      | 63        |
| LoSim-HiSuc | 15.70    | 3.43      |          |           | 2.79     | .49       |          |           |
| HiSim-LoSuc | 18.85    | 4.58      | 3.83**   | 50        | 2.77     | .50       | 1.71*    | 76        |
| LoSim-LoSuc | 14.38    | 3.77      |          |           | 2.56     | .56       |          |           |
| LoSim-HiSuc | 15.70    | 3.43      | 1.27     | 47        | 2.79     | .49       | 1.79*    | 71        |
| LoSim-LoSuc | 14.38    | 3.77      |          |           | 2.56     | .56       |          |           |

\*  $p < .05$   
\*\*  $p < .01$

Results from analyses using the composite ACT scores were more ambiguous. Examination of the relevant values in Table 2 reveal only partial support for the processing model. Three predictions were supported; three were not. The mean composite ACT for the HiSim-HiSuc group was significantly higher than the mean scores obtained by the LoSim-HiSuc group and the LoSim-LoSuc group. Also, the HiSim-LoSuc group mean score was significantly higher than the mean obtained by the LoSim-LoSuc group. However, contrary to prediction, the mean HiSim-HiSuc



group score was not significantly greater than the mean from the HiSim-LoSuc group. Nor was the mean from the LoSim-HiSuc group significantly higher than the mean from the LoSim-LoSuc group. Finally, the two intermediate group means were not equivalent; the HiSim-LoSuc group mean was significantly higher than the mean from the LoSim-HiSuc group.

### *Discussion*

The strong two-factor solution obtained from the initial factor analysis of the six cognitive tasks was anticipated. The solution is consistent with a rational analysis of task demands and with findings from numerous studies (see Das et al., 1979). Since these factors did emerge, further analyses were conducted to investigate the relationship of college achievement to simultaneous-successive processing.

When cumulative GPA was employed as the dependent measure, all predictions were confirmed. The cumulative GPA reflects performance in a wide variety of skill areas; apparently both simultaneous and successive processing substantially mediate broadly-based classroom performance. Individuals within the HiHi group performed better than those high in only one area; those high in only one processing dimension out-performed those in the LoLo group; and those in the two "off-diagonal" groups performed equally well. These results provide support for the model and are consistent with the contention that the modes are nonhierarchical.

Three of the six predictions using the composite ACT score were not supported—the HiSim-HiSuc group did not earn a higher mean ACT score than the HiSim-LoSuc group; the LoSim-HiSuc group did not earn a higher mean ACT score than the LoSim-LoSuc group; and, finally, the HiSim-LoSuc group did earn a significantly higher mean score than the LoSim-HiSuc group. Thus, the contrasts not conforming to predictions substantiate the relative importance of the simultaneous processing mode for attaining higher ACT scores. But because test reviewers (e.g., Wallace, 1975) have criticized the ACT for a heavy reading component, the exceptions may be interpreted within the context of previous results which emphasize the importance of simultaneous processing skills for complex reading (Cummins & Das, 1977). That is, if the ACT is interpreted as a measure of sophisticated reading skills, then the superordinate status of simultaneous processing in the ACT measure may be explained by the theory.

In summary, the present results are consistent with the Luria-Das model. Specifically, there appears to be a relationship between cognitive processing ability and academic achievement of college students. In general, higher simultaneous processing is important for superior ACT performance whereas higher simultaneous and successive processing are necessary for superior GPA.

### *Implications*

Rigorous experimental hypotheses testing of the Luria-Das model is clearly needed. Thus far, educators have been relatively unsuccessful in identifying appropriate treatment x aptitude models (Reynolds, 1981). But research into cognitive functioning within this model appears promising. For example, once simultaneous or successive deficits have been identified, remedial strategies become apparent. Means of identifying simultaneous-successive deficits include the use of the battery and methodology described within the present paper. A second method, currently becoming available, is called the Kaufman Assessment Battery for Children (K-ABC) (Kaufman & Kaufman, 1983). The availability of this instrument and the in-

creasingly abundant research findings investigating the model should hasten efforts to explore the efficacy of the model within educational and clinical settings.

#### Notes

1. The model-based hypotheses of the present study were based on the assumption that the dependent achievement measures required both processing modes in about equal proportions. Therefore, the use of multiple a priori *t*'s was justified, and the concomitant increased probability of Type I error was tolerated. In retrospect, this assumption was not justified for the ACT dependent measure (i.e., simultaneous processing is apparently more important), but the *t* ratios of the ACT measure exceed the more conservative  $\alpha < .01$  value in all cases of significance.

#### References

- Beery, K. E., & Buktenica, N. A. (1967). *Developmental test of visual motor integration: Student test booklet*. Chicago: Follett.
- Bender, L. (1946). *Visual motor gestalt test*. New York: American Orthopsychiatric Association.
- Cummins, J., & Das, J. P. (1977). Cognitive processing and reading difficulties: A framework for research. *Alberta Journal of Educational Research*, 23, 245-256.
- Das, J. P., & Jarman, R. F. (1980). Coding and planning processes. In M. Friedman, J. P. Das, & N. O'Connor (Eds.), *Intelligence and learning*. New York: Plenum Press.
- Das, J. P., Kirby, J., & Jarman, R. F. (1975). Simultaneous and successive syntheses: An alternative model for cognitive abilities. *Psychological Bulletin*, 82(1), 87-103.
- Das, J. P., Kirby, J. R., & Jarman, R. F. (1979). *Simultaneous and successive cognitive processing*. New York: Academic Press.
- Dolch, E. C. (1942). *Basic sight-word test*. Champaign, IL: Garrard.
- Jarman, R. F. (1978). Level I and Level II abilities: Some theoretical reinterpretations. *British Journal of Psychology*, 69, 257-269.
- Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman assessment battery for children*. Circle Pines, MN: American Guidance Service.
- Kaufman, P., & Kaufman, D. (1979). Strategy training and remedial techniques. *Journal of Learning Disabilities*, 12, 63-66.
- Krywaniuk, L. W., & Das, J. P. (1976). Cognitive strategies in native children: Analysis and intervention. *Alberta Journal of Educational Research*, 22, 271-280.
- Luria, A. R. (1966a). *Higher cortical functions in man*. New York: Basic.
- Luria, A. R. (1966b). *Human brain and psychological processes*. New York: Harper & Row.
- McCallum, R. S., & Merritt, F. M. (1983). Simultaneous-successive processing among college students. *Journal of Psychoeducational Assessment*, 1, 85-93.
- Merritt, F. M., & McCallum, R. S. (1983). Sex-related differences in simultaneous-successive information processing. *Clinical Neuropsychology*, 5, 117-119.
- Raven, J. C. (1958). *Standard progressive matrices: Sets A, B, C, D, and E*. London: H. K. Lewis.
- Reynolds, C. R. (1981). Neuropsychological assessment and the habilitation of learning: Consideration in the search for the aptitude x treatment interaction. *School Psychology Review*, 10, 343-349.
- Slosson, R. L. (1963). *Slosson oral reading test*. New York: Slosson Educational Publications.
- Wallace, W. L. (1975). Review of the ACT test battery of the American College Testing Program. In O. K. Buros (Ed.), *Intelligence tests and reviews*. Highland Park, NJ: Gryphon Press.

## ESSAY REVIEWS

### LOGO AND THE DEVELOPMENT OF GENERAL PROBLEM-SOLVING SKILLS

Linda R. Krasnor  
and

John O. Mitterer  
*Brock University*

How to structure children's encounters with computers and what to expect from those encounters are questions with far-reaching educational implications. Particularly high expectations have been aroused by LOGO, a graphics-oriented structured language which is imbedded in an educational philosophy stressing play, discovery and mastery (Papert, 1980; Groen, 1978). LOGO allows children to direct a "turtle" around a video screen to create designs. Once a small set of primitives has been learned, children can "teach" the computer to draw designs by defining procedures. Each child can develop a set of procedures which can be combined, just like the primitives, into higher order procedures to make complex designs. Thus the child gradually builds a personal computer language while learning fundamental programming concepts.

Papert (1980) has argued that when children experience LOGO they learn much more than a computer language. The LOGO experience was designed explicitly to facilitate the learning of powerful ideas, skills and heuristics which transcend the immediate task environment and can be applied in other problem-solving situations. This is an exciting possibility since curriculum designers have traditionally focused on specific content rather than general problem-solving skills (Brown, Bransford, Ferrara, & Campione, 1983; Larkin, 1980).

It is not implausible that some skills learned in programming may transfer to other types of problem-solving since it has often been noted that computer programming *is* a type of problem-solving. The steps in designing a program closely

---

Linda Rose Krasnor is a Developmental Psychologist with research interests in social interaction during problem solving activities. She is at Brock University.

John Mitterer is a Cognitive Psychologist interested in cognitive development. Recently, he has become interested in the applications of cognitive and developmental theories to education.



parallel generally recognized steps in the problem-solving process (Statz, 1973; Chait, 1978; Papert, 1980). In Polya's (1957) classic model of problem-solving, four basic steps were outlined: understanding the problem, devising a solution, carrying out the solution plan, and evaluating the solution. Similarly, the programmer must understand the programming task and be able to represent the problem. The program must be devised, entered into the computer, and debugged.

Several skills have consistently been advanced as powerful ideas which can be acquired through the LOGO experience (Papert, 1980). One is the strategy of breaking problems into manageable subproblems. Another is means-end analysis or the systematic planning of actions to achieve goals (e.g., Sacerdoti, 1977). A third is debugging, the principle that problem solutions can be successively refined. The development of a more positive attitude to errors is a fourth powerful idea. Programming bugs are recognized as being unavoidable and useful in providing information. Such a reconceptualization of failure implies that the child's sense of mastery and motivation would not be negatively affected by errors.

In this paper, we will provide a framework for assessing the extent to which these skills learned through LOGO may be transferred to other problem-solving contexts. Focusing on the turtle geometry aspects of LOGO, we will examine some of the principles governing the generalization of learning as they apply to the LOGO experience and will consider the existing empirical evidence for the transfer of these skills to other domains. We will also consider how transfer may depend on individual differences among children. Finally, we will isolate specific components of the complex LOGO experience in an attempt to specify which components may be responsible for transfer if and when it does occur.

### *The Generalization of Learning*

As Simon (1980) has made clear, the notion of general knowledge is really the notion that some kinds of knowledge can be transferred from the context of acquisition to a wide variety of other situations. Thus, a consideration of the circumstances which will facilitate the transfer of learning from one situation to another will in effect be a consideration of the factors governing the generality of knowledge. From this perspective, recent research has most often failed to produce evidence for the transfer of knowledge between domains (see Brown, et al., 1983). It now appears that much domain specific, nontransferable knowledge is required for expertise in any domain. As Simon (1980) has suggested, "No one, no matter how intelligent, skilled in problem solving, or talented, becomes a chess grandmaster without 10 years of intense exposure to the task environment of chess" (p. 82).

While it is possible that the basic strategy of training general skills for transfer may not be as effective as previously supposed, there are, nevertheless, many factors which have long been recognized to influence transfer (Ellis, 1965). One important principle is that transfer between tasks requires that some of the processes or knowledge used in the tasks be essentially identical (Thorndike & Woodworth, 1901). A second is that the learner must recognize that a new problem situation is similar to one encountered previously (e.g., Woodworth & Schlosberg, 1954). It has commonly been found that little spontaneous transfer occurs between well-formed problem isomorphs (e.g., Reed, Ernst, & Banerji, 1974; Hayes & Simon, 1977). If explicit instruction is given in the range of applicability of a general principle (Brown, 1978; Brown & Campione, 1981), however, or if learners are encouraged to state a general rule (Gagne & Smith, 1962; Thorndyke & Stasz, 1980), then transfer may be

improved. Third, the degree of transfer may depend on the completeness of original learning (Brown et al., in press). Finally, the degree of transfer may depend on exposure to a variety of situations in which the specific skill is useful. Training in multiple contexts has been found to improve transfer, when compared to relatively uniform learning contexts (Belmont, Butterfield, & Borkowski, 1978; Brown, 1974, 1978; Brown, Campione, & Day, 1981).

Therefore, transfer (if it occurs at all) would be expected between highly similar domains under circumstances where learners were fluent with the skills in question, had practised those skills in several situations, and were explicitly aware of the general utility of the skills. With these considerations in mind, to what extent would the four powerful ideas from LOGO be expected to transfer to other problem domains?

The only condition for transfer to be consistently met by such an experience is that LOGO problems are similar to problems in other domains. However, this similarity may well be at a relatively abstract level and many students will need explicit guidance in order to recognize correspondences. LOGO instructional experiences typically have not included any attempt to explicitly teach students in the use of these skills. Rather, tutors rely on subtle hints and structuring of problems to achieve this end. This may mitigate against transfer rather than facilitating it.

LOGO instruction is generally short-term; a typical LOGO experience consists of 20 to 40 hours of hands-on experience. There is thus little opportunity to ensure sufficient mastery of programming fundamentals to successfully apply means-end and sub-goal analysis to many different problems. Also, it is quite possible that the original learning has taken place in a very narrow context, especially for those children who work only with turtle graphics. The lack of practice across a variety of situations may reduce the opportunity for transfer even further.

The situation is somewhat improved with respect to the powerful ideas of debugging and a positive attitude towards bugs. In LOGO instruction, tutors are advised to encourage exploration after errors and to minimize the negative consequences of errors. Children are exposed to debugging very early in the instructional sequence, since they make errors as soon as they begin. Students thus have the opportunity to edit procedures repeatedly. Considerable expertise with these skills should be built up. Lack of explicit instruction in their general applicability and a lack of practice in different contexts may reduce the degree of transfer, however, even for these powerful ideas.

On balance, we simply cannot predict how much transfer of these powerful ideas to expect from the LOGO experience. However, we can observe that some of the most important factors in promoting the transfer of other skills are not present in the LOGO experience as it is most commonly described in the literature. Clearly, there are many research opportunities presented by the questions raised here. As the following review of the literature will reveal, few researchers have examined these important issues.

### *Evidence for the Generality of LOGO Learning*

To date the literature concerning LOGO consists primarily of testimonials (e.g., Goldenberg, 1980; Watt, 1982), curriculum guides (e.g., Birch, 1980), manuals (e.g., Abelson, 1982) and a few studies containing considerable anecdotal evidence. Although researchers have explored the use of LOGO with adults (Brown & Rubenstein, 1974; Lukas & Feurzeig, 1972; Austin, 1976; duBoulay, 1977), children



have received the most attention. Most of the available evidence refers to LOGO learning rather than the transfer of LOGO learning to other contexts. This body of work has provided anecdotal evidence that the powerful ideas mentioned above can be learned by children and a suggestion that these skills can be transferred to other domains.

The work of the M.I.T. LOGO group has been typical of this research approach. These researchers have focused on case studies (e.g., Lawler, 1980) and the creation of curriculum materials (e.g., Bamberger, 1979) rather than experimental findings. The Brookline project is a representative example. This collaboration between the M.I.T. LOGO group and the public schools of Brookline, Massachusetts, resulted in detailed anecdotal descriptions of the learning of 16 children without systematic experimentation or statistical analysis (see Papert, Watt, diSessa, & Weir, 1979).

A few quotes from the interim report of this project (Papert, Abelson, Bamberger, diSessa, Weir, Watt, Hein, & Dunning, 1978) will suffice to illustrate these descriptive data. The comments concern a single child, Ray. The first two were made by one of the authors (Watt) who noted,

From the comments made by his teachers, and from observations made by myself and others in the LOGO classes, I have a model of Ray as a boy who has chosen to react to academic difficulties by adopting a pose of indifference, and refusing to take responsibility for his work or behavior in school. (Appendix 1, Individual profiles; Ray, p. 1)

Watt goes on, "to *speculate* that from his earliest school experiences, Ray has been afraid of failure, and especially of *appearing* to fail" (p. 2) (author's italics). Finally, after learning LOGO, one of Ray's teachers commented, "The breakthrough for him in LOGO, the success he has had, is powerful information for me . . . he has produced the best piece of writing I've seen from him" (p. 12). Such descriptions definitely suggest the transfer of important abilities from LOGO to another domain, but very little objective analytic evidence is provided.

This research approach is consistent with both published (Papert, 1980) and private (Solomon, 1982, personal communication) claims that traditional experimental research methodologies are inappropriate for documenting the effects of LOGO. Much of the other published work on LOGO suffers from a similar lack of empirical documentation (e.g., Feurzeig & Lukas, 1972; Weir & Emanuel, 1976). This view has several sources: (a) that traditional psychological and educational testing methods are not appropriate for measuring LOGO-facilitated changes; (b) that group-based designs and statistical analyses will obscure the wide individual differences in children's learning patterns and adult teaching styles; and (c) that standardization of LOGO teaching procedures will substantially detract from the flexibility and individualization of the LOGO experience.

Only a handful of studies have produced empirical evidence relevant to the claim that learning LOGO produces general effects. Of these, the Edinburgh LOGO project (e.g., Howe, O'Shea, & Plane, 1979) focused on the development of mathematical skills and has produced mixed evidence of benefits while the thesis work of Statz (1973) and Chait (1978) focused on the development of general problem-solving skills.

Of the two reports, that of Statz (1973; see also Statz, Folk, & Seidman, 1973) more closely approximated a classical experimental design. Problem-solving was investigated by administering pre- and posttests to a group of 16 grade 4 children who had participated in LOGO sessions and to a similar group who remained in the normal class environment. The two groups were administered four problem-solving



tests. The Twenty Questions Game (Bruner, Rose, Oliver, & Greenfield, 1967) was used to investigate the child's ability to organize concepts hierarchically, as revealed through constraint-seeking questions. The Tower of Hanoi puzzle (e.g., Simon, 1975) was used to test the ability to develop a systematic solution strategy. To assess permutation abilities, the children were also asked to generate all possible orders of sets of numbers. Finally, a word task was used to assess classification skills. Statz (1973) predicted that learning LOGO would lead to improved performance on all of these problem-solving measures. While she found trends in the anticipated direction for all of the tasks, only the permutation and the word puzzle tasks reached significance.

Chait (1978) investigated comparable issues with five grade 7 children who were taught individually. These children were administered pre- and posttests but no control group was employed. The test battery included a test of basic arithmetic skills, an alternate uses test, and a set of geoboard problems including duplications, rotations, and mirror-imaging. The results revealed no significant changes in test scores.

The entire LOGO literature merits considerable criticism from an experimental point of view. The general lack of objective measurement makes the currently available work useful only as a first step in evaluating the effects of LOGO. The work of Statz reveals several problems which limit the usefulness of her work. First, the training procedure she used was modified partway through the study, making assessment difficult. Second, her measures of problem-solving were not tightly tied to any extant approach to problem-solving (e.g., Polya, 1957; Newell & Simon, 1972), although they might be considered to reflect the ability to carry out means-end analysis. The work of Chait is similarly open to criticism. A small sample size and the lack of a control group would have made the results difficult to interpret even had they reached significance. The test battery was also not tied to any theory of problem-solving. It is not clear why any of the tasks she chose should reveal any change as a result of learning LOGO.

To summarize, there is as yet no good evidence that any of the powerful ideas mentioned above generalize to other domains. This is either because existing research reports do not convincingly test for generalization (as in the case of means-end analysis) or because the relevant research has not yet been done. Clearly, in light of the potential impact of LOGO, these issues cannot remain unaddressed, and their resolution promises to be a fruitful area for future investigation.

### *Individual Differences and the Generality of LOGO Learning*

One of the striking aspects of the current literature has been the extensive range of applicability claimed for LOGO. Learning LOGO is presumed to benefit normal children as young as three or four years of age. While there is considerable anecdotal evidence for the effectiveness of LOGO in teaching programming to such young children, more controlled research has indicated that preschoolers do not have the necessary spatial concepts to effectively use the turtle's intrinsic geometry and are unable to connect specific keyboard activity to robot movement (Gregg, 1978).

Children who are unable to utilize the intrinsic geometry system or to understand the relation between program commands and turtle movement would not achieve the mastery and control objectives of the LOGO experience. Children with inadequate language, spatial, and mathematical skills would have limited possibili-

ties for going beyond immediate programming modes. While it is likely that elementary computer literacy could be achieved, relatively little generalizable problem-solving skill would be expected without learning structured programming.

Young and relatively poor learners may need to be taught both specific *and* general skills in a given content domain in order to show transfer (Brown et al., 1983). More able learners may only need to be taught more specific skills, since they are more likely to possess general heuristics already. More advanced learners may also spontaneously prepare for transfer by doing such things as hypothesizing applications and using analogies. This suggests that younger children need to be explicitly taught general skills, as well as skills specific to the turtle world, before transfer could be reasonably expected. The discovery-oriented nature of the LOGO experience does not provide this explicitness for general skills. Thus, only older students may be able to apply LOGO skills to other domains.

Generalization of problem-solving skills may well also depend on the child's Piagetian developmental level. On one hand, Papert (1980) has argued that LOGO may offer to the child a way of concretizing the abstract, a very powerful tool for the attainment of formal operational thought. On the other hand, it is quite possible that only children who have reached the formal operational stage will be capable of fully utilizing LOGO (e.g., Folk, 1972). Both positions suggest that an optimal age for LOGO instruction may be the onset of the formal operational stage—about 12 years of age (Inhelder & Piaget, 1958). It is reasonable to assume that some general problem-solving heuristics will require formal operational thought. A child at a pre-operational level would be unlikely to achieve the hierarchical thinking necessary for structured programming. Such thinking may well be required to master subgoal analysis and mean-end planning. Even concrete operational children would likely find it difficult to manipulate symbols logically and to grasp concepts of recursion, multiple inputs, and permutation at an abstract level. Yet it is at this more advanced level that generalization of these skills would most likely occur. In contrast, it is likely that much younger children would be capable of understanding and benefiting from the powerful ideas of debugging and adaptive response to errors since it is less obvious how these skills may depend upon a foundation of formal operational thought. Although this is an issue of theoretical and applied import, to date there has been no systematic investigation of the relationship between developmental level and the generalization of problem-solving skills learned through LOGO.

In an early investigation, Folk (1972) has directly examined the possibility that the ability to learn LOGO fully in the first place may depend on the developmental level of the child. Unfortunately, Folk's unpublished thesis provides little more than a suggestive set of hypotheses. Thus, the extent of LOGO learning which may be expected for a child of a specific age or developmental level is still an open question. It is possible that only formal operational children will be able to learn LOGO fully and transfer LOGO ideas such as means-end analysis and subgoal analysis to other contexts. Adaptive response to failure, however, may be learned and generalized by much younger children.

In addition to a wide range of typical children, there has been a keen interest in the use of LOGO with special groups of children such as autistic, educable retarded, and dyslexic children (e.g., Goldenberg, 1980). The appeal of LOGO for these hard-to-treat groups is understandable. LOGO's novelty and immediate appeal, however, cannot ensure the sustained interest necessary to develop general problem-solving



skills. Without an understanding of the potential of LOGO with typical children, its use with atypical children may be built on relatively uncertain ground. It is likely that such special groups of atypical children would be much less likely to develop general problem-solving skills than normal children. Clearly, this issue provides much scope for future research.

### *Identifying the Components of LOGO Which Produce General Learning*

A detailed task analysis of the LOGO experience will be required before the problem of generalization can be completely specified. As a first step, we have isolated several educationally potent components of the experience, each of which may be sufficient to induce cognitive and behavioral change. Two of these components are aspects of the LOGO language itself, two are aspects of the accompanying educational philosophy, and the final one is the motivating character of LOGO. The combination of these diverse factors in the LOGO experience makes it difficult to assess the unique contribution of LOGO *as a language* to learning and transfer.

The first feature of the language is the turtle, which is intended to provide the child with a way to "concretize the abstract" (Papert, 1980). By encouraging the child to "play turtle," syntonic learning is made possible. Associated with the turtle is intrinsic geometry or turtle geometry, a non-Cartesian approach to geometry (Abelson & diSessa, 1981). In turtle geometry, the state of the turtle is specified without respect to an external reference frame. This makes it easier to begin graphics programming since only the commands, LEFT, RIGHT, FORWARD, and BACK are required to specify turtle state, commands which derive meaning from the child's own physical experience. In addition, graphics-based programming environments may have a great appeal due to the potency of visual experience. The use of such a concrete medium may facilitate the ability to engage in means-end and subgoal analysis by making the task of completely conceptualizing a problem much simpler.

The second feature of LOGO is the ease with which structured programming is possible. New procedures can be created at will and freely combined with other procedures to create complex procedures more easily than in any other available language. In a compiled language like PASCAL, for example, a listing of the entire program (including all subprocedures) must be brought together for compilation. This is not necessary in LOGO, in which defined procedures can be treated just like primitives in developing more complex procedures. Subgoal and means-end analysis are clearly facilitated by such capabilities.

These two components of LOGO jointly create a programming environment characterized by immediate feedback and total control. As a LOGO graphics program is executed, each step of the program is displayed. Most bugs stand out and may be debugged right away. This should facilitate the acquisition of the idea that debugging is a useful problem-solving strategy. Compiled, text-oriented programming may not allow such efficient feedback. Such an environment, which is responsive to attempts at control, may allow children to develop a sense of mastery and competence (Bronson, 1974). In fact, it is one of the major claims of LOGO proponents (e.g., Papert, 1980) that the microworld of the turtle is one in which the child can feel powerful and successful.

In addition to these two components, several other elements of the LOGO experience can be identified which are not inherently tied to the language itself. One such component is a stress on play and exploration; the second is the small group nature of the LOGO experience.



In the LOGO environment, the child is free to experiment without risking any real-life consequences. This atmosphere of play and exploration is aided by allowing the child to select projects, thus creating a personal involvement and commitment which may nurture greater success. The child's affect is almost always described as positive, and the turtle world is certainly nonliteral. As such, the LOGO experience satisfies most major definitions of play (Krasnor & Pepler, 1980). The child is encouraged to try out new means of achieving goals and to combine old means to achieve new goals. This may result in a more flexible and adaptive mastery of LOGO than a more regimented approach, since it has been shown that free play with materials can lead to enhanced creativity and problem-solving (e.g., Dansky & Silverman, 1973, 1975; Smith & Dutton, 1979; Sylva, Bruner, & Genova, 1976). Thus it is likely that children who learn LOGO in an exploratory, play-based mode will be more adept at solving programming problems than children who are taught LOGO in a more directive form. In addition, adaptive attitudes toward failure and a willingness to debug problem solutions should be facilitated.

In most implementations of the LOGO curriculum, children share equipment and are encouraged to cooperate with peers to complete small group projects. This social arrangement can be considered to be another major component of the LOGO environment. Group problem-solving can be an intensive experience in itself, and may be an important factor in any changes which have previously been attributed to LOGO. The child has the opportunity to learn by observing the nature and outcome of others' actions (Bandura, 1977). In addition to benefiting from peer suggestions, the child may also engage in teaching others. This tutoring may help the child to consolidate and practise skills (Allen, 1976). Further, there is a need for the child to communicate clearly and effectively with peers, both in order to teach and to negotiate joint programming projects. Practice in these skills may lead to improvement in communication ability.

The negotiation and confrontation present in social interaction is thought to be an important factor in cognitive development, especially in the growth of decentration and perspective-taking skills (Piaget & Inhelder, 1969). The effects of small-group project-oriented interactions thus may be powerful ones, and there has been no control for its influence in prior research. Thus it is quite possible that children who learn LOGO in a small group will become more adept at programming in LOGO and at referential communication than children who learn LOGO alone. In addition, increased peer social status may result from attaining expertise in a new skill, and by the formation of new friendships through shared cooperative experience.

Finally, one of the assumed advantages of learning problem-solving skills through LOGO programming is that it is fun. This motivation may be a critical component of success in teaching LOGO and in producing any resulting transfer. A distinction should be made, however, between motivation for initially establishing contact with the LOGO experience, and the type of motivation which is needed for sustaining interest in LOGO over extended periods of time. There is little question that microcomputers and turtle graphics will provide sufficient novelty to motivate initial contact with the language. This initial motivation must carry the child through the period necessary to learn the basic fundamentals of the language and its technology. For a young child, this is not a trivial accomplishment. The child must learn the keyboard symbols, understand simple editing, recognize the need for exact syntax and supplying both direction and quantitative information. If the child does not continue efforts through this technique period, mastery of the skills neces-

sary to implement programming ideas and successfully solve problems will not develop.

To summarize, it is quite possible that the powerful ideas of debugging and positive response to failure can be transmitted simply by exposure to play-oriented, small-group experiences, even without any LOGO programming. However, the other powerful ideas of means-end analysis and subgoal analysis seem more tightly tied to the language itself. Even here, though, it is quite possible that noncomputer problem-solving curricula stressing these heuristics can be just as successful as LOGO training. It is in the motivational component, perhaps, that LOGO has a special advantage over other methods of teaching problem-solving. Alternately, it is possible that the overall instructional package offered by the LOGO experience, blending as it does so many potent instructional devices into one curriculum, provides by far the best context for facilitating the learning and transfer of these powerful ideas. Only the systematic investigation of the contribution of these and other components of the LOGO experience will result in definitive answers.

### *Conclusions*

In our review of current research, we have found no evidence that the LOGO experience with turtle geometry will facilitate the development of general problem-solving skills. Further, there is some theoretical reason for doubting that such skills will be developed by the LOGO experience. Until further research resolves this question, educators interested in using LOGO in this way should tailor the LOGO environment to facilitate transfer. In general, the effective use of LOGO to develop problem-solving skills likely depends on three central conditions: (a) the use of techniques which foster awareness of the general utility of problem-solving heuristics; (b) selection of an appropriate group of children; and (c) the inclusion of the major components of the LOGO environment, since it is possible that LOGO programming per se will not be sufficient to produce transfer.

Tutors should explicitly emphasize how the skills learned in LOGO are relevant for problems outside the turtle world, how children may recognize appropriate similarities among problem types, and how LOGO skills might be adapted for other types of problems. Children should be encouraged to solve a variety of graphic and nongraphic problems, both in and out of the LOGO environment. Furthermore, it is likely that longer exposures to LOGO will increase mastery of general programming skills, which may enhance transfer. Extended training periods may be particularly critical for weak students, who will need more time to master the specific content of LOGO and to learn the general transfer skills. Much more research will be required to specify a curriculum using LOGO to promote general problem-solving skills.

On the basis of current evidence, it appears that only formal operational children of at least 12 years of age will be able to benefit fully from attempts to foster general problem-solving through LOGO. For younger or less able children, expectations for LOGO should be focused on mastery, adaptive reaction to failure, and computer literacy.

Attempts to fit LOGO into more traditional curricula (e.g., by grading projects, or predetermining the sequence of concepts learned) may reduce its potential benefit. If this is done, children may not have adequate opportunity to explore and play with LOGO, although play and exploration may be the most critical component for developing powerful ideas. Similarly, isolating the child with a single computer or otherwise regulating peer interaction in the interests of classroom management may limit the motivational power of LOGO and the benefits of group problem-solving.



Although we have focused upon LOGO's possibilities for enhancing general problem-solving and have attempted to highlight the potential limitations of LOGO for this purpose, there are many other reasons why children of a wide range of abilities and developmental levels should have an opportunity to experience LOGO. First, LOGO is fascinating in its own right. Further, LOGO may be a superior language for attaining computer literacy and introducing programming fundamentals, especially for younger children. Future research contrasting LOGO with other languages such as BASIC will be interesting. Finally, further research may show that the LOGO experience promotes the attainment of other educational objectives, such as fostering self-esteem, communication abilities, and a sense of mastery.

So far, there has been a tendency to overlook the need for objective and systematic examination of the LOGO experience and its effects. Without this appraisal, there is a distinct possibility that exciting learning opportunities may be lost. In the enthusiasm surrounding the availability of microcomputers and the development of languages such as LOGO, educators have envisioned dramatic and pervasive benefits for children. These expectations, if untempered, may eventually work against the productive use of computers in education. Unrealistic beliefs concerning the benefits of LOGO may result in frustration and disillusionment. Without analytic empirical investigations of LOGO, educators will lack the knowledge needed to design efficient and effective ways of using LOGO.

Our traditional knowledge of the principles of learning and transfer are still relevant, in spite of the futuristic nature of the educational technology. Past research, some of which dates to the beginning of this century, provides an important foundation for structuring LOGO learning environments and for predicting realistic outcomes. This blending of the old and the new provides an exciting and possibly unique research opportunity.

The authors especially wish to thank Joan Preston for her early participation on our LOGO project. In addition, Cynthia Solomon, Sharnée Chait (and LCSJ), Allison Birch (and the Phoenix School), Sandra Wills, Dave Darvill, and Dale Burnett provided many stimulating thoughts. The preparation of this article was funded in part by Social Sciences and Humanities Research Council of Canada strategic seed grant #499-82-1010.

### References

- Abelson, H. (1982). *LOGO programming for the Apple II*. Peterborough, NH: Byte/McGraw Hill.
- Abelson, H., & diSessa, A. (1981). *Turtle geometry: The computer as a medium for exploring mathematics*. Cambridge: M.I.T. Press.
- Allen, V. (1976). Children helping children: Psychological processes in tutoring. In J. Levin & V. Allen (Eds.), *Cognitive learning in children: Theories and strategies*. New York: Academic.
- Austin, H. (1976). Teaching teachers LOGO: The Lesley experiments. *M.I.T. LOGO Memo #23*.
- Bamberger, J. (1979). LOGO music projects: Experiments in musical perception and design. *M.I.T. LOGO Memo #52*.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Belmont, J., Butterfield, E., & Borkowski, J. (1978). Training retarded people to generalize memorization methods across memory tasks. In M. Gruneberg, P. Morris, & R. Sykes (Eds.), *Practical aspects of memory*. London: Academic Press.



- Birch, A. (1980). *LOGO curriculum for the Lincoln-Sudbury Schools, Mass.* Unpublished manuscript.
- Bronson, W. (1974). The growth of competence: Issues of conceptualization and measurement. In K. Connolly & J. Bruner (Eds.), *The growth of competence*. London: Academic Press.
- Brown, A. (1974). The role of strategic behavior in retardate memory. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 7). New York: Academic Press.
- Brown, A. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 1). Hillsdale, NJ: Erlbaum.
- Brown, A., Bransford, J., Ferrara, R., & Campione, J. (1983). Learning, remembering and understanding. In J. Flavell & E. Markman (Eds.), *Carmichael's manual of child psychology* (Vol. 1). New York: Wiley.
- Brown, A., & Campione, J. (1981). Inducing flexible thinking: A problem of access. In M. Friedman, J. Das, & N. O'Connor (Eds.), *Intelligence and learning*. New York: Plenum.
- Brown, A., Campione, J., & Day, J. (1981). Learning to learn: On training students to learn from texts. *Educational Researcher*, 10, 14-21.
- Brown, J. S., & Rubenstein, R. (1974). *Recursive functional programming for the student in the humanities and social sciences* (Technical Report No. 27a). Department of Information and Computer Science, University of California.
- Bruner, J. (1972). The nature and uses of immaturity. *American Psychologist*, 27, 687-708.
- Bruner, J., Rose, R., Oliver, P., & Greenfield, P. (1967). *Studies in cognitive growth*. New York: Wiley.
- Chait, S. (1978). An analysis of children's problem solving in a graphics oriented computer programming environment. Unpublished master's thesis, McGill University, Montreal, Quebec.
- Dansky, J., & Silverman, I. (1973). Effects of play on associative fluency in preschool-aged children. *Developmental Psychology*, 9, 38-49.
- Dansky, J., & Silverman, I. (1975). Play: A general facilitator of associative fluency. *Developmental Psychology*, 11, 104.
- duBoulay, B. (1977, March). Learning teaching mathematics. *Mathematical Teaching*, 52-57.
- Ellis, H. (1965). *The transfer of learning*. New York: Macmillan.
- Feurzeig, W., & Lukas, G. (1972). LOGO: A programming language for teaching mathematics. *Educational Technology*, 12, 39-46.
- Folk, M. (1972). Influences of developmental level on a child's ability to learn concepts of computer programming. Unpublished doctoral dissertation, Syracuse University, Syracuse, NY.
- Gagne, R., & Smith, E. (1962). A study of the effects of verbalization in problem-solving. *Journal of Experimental Psychology*, 63, 12-18.
- Goldenberg, P. (1980). *Special technology for special children*. Cambridge, MA: M.I.T. Press.
- Gregg, L. (1978). Spatial concepts, spatial names, and the development of exocentric representations. In R. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, NJ: Erlbaum.
- Groen, G. (1978). The theoretical ideas of Piaget and educational practice. In P. Suppes (Ed.), *Impact of research in education: Some case studies*. Washington, DC: National Academy of Education.
- Hayes, J., & Simon, H. (1977). Psychological differences among problem isomorphs. In N. Castellan, J. D. Pisoni, & G. Potts (Eds.), *Cognitive theory* (Vol. 2). Hillsdale, NJ: Erlbaum.
- Howe, J., O'Shea, T., & Plane, F. (1979). Teaching mathematics through LOGO programming: An evaluation study. *DAI Research Paper No. 115*.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic.

- Krasnor, L., & Pepler, D. (1980). The study of children's play: Some suggested future directions. In K. Rubin (Ed.), *Children's play: New directions for child development* (Vol. 9). San Francisco: Jossey-Bass.
- Larkin, J. H. (1980). Teaching problem solving in Physics: The psychological laboratory and the practical classroom. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research*. Hillsdale, NJ: Erlbaum.
- Lawler, R. (1980). The progressive construction of mind (One child's learning: addition). *M.I.T. LOGO Memo #57*.
- Lukas, G., & Feurzeig, W. (1972, October). Computers and learning to learn. *Proceedings of the second City University of New York Conference, Innovations in Educational Technology*.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Papert, S. (1980). *Mindstorms: Children, computers and powerful ideas*. New York: Basic.
- Papert, S., Abelson, H., Bamberger, J., diSessa, A., Weir, S., Watt, D., Hein, G., & Dunning, S. (1978). Interim report of the LOGO project in the Brookline Public Schools. *M.I.T. LOGO Memo #49*.
- Papert, S., Watt, D., diSessa, A., & Weir, S. (1979). Final report of the Brookline LOGO project. Part II: Project summary and data analysis. *M.I.T. LOGO Memo #53*.
- Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic.
- Polya, G. (1957). *How to solve it*. New York: Anchor.
- Reed, S., Ernst, C., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology*, 6, 436-450.
- Sacerdoti, E. (1977). *A structure for plans and behavior*. New York: Elsevier North-Holland.
- Simon, H. (1975). The functional equivalence of problem-solving skills. *Cognitive Psychology*, 7, 268-288.
- Simon, H. (1980). Problem solving in education. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research*. Hillsdale, NJ: Erlbaum.
- Smith, P., & Dutton, S. (1979). Play and training on direct and innovative problem solving. *Child Development*, 50, 830-836.
- Statz, J. (1973). *The development of computer programming concepts and problem solving abilities among ten-year-olds learning LOGO*. Unpublished doctoral dissertation, Syracuse University, Syracuse, New York.
- Statz, J., Folk, M., & Seidman, R. (1973). *Final report: Syracuse LOGO project*.
- Sylva, K., Bruner, J., & Genova, P. (1976). The role of play in the problem-solving of children 3- to 5-years old. In J. Bruner, A. Jolley, & K. Sylva (Eds.), *Play*. New York: Basic.
- Thorndike, E., & Woodworth, R. (1901). I. The influence of improvement in one mental function upon the efficiency of other functions; II. The estimation of magnitudes; III. Functions involving attention, observation and discrimination. *Psychological Review*, 8, 247-261, 384-395, 553-564; cited in Ellis (1965).
- Thorndyke, P., & Stasz, C. (1980). Individual differences in procedures for knowledge acquisition from maps. *Cognitive Psychology*, 12, 137-175.
- Weir, S., & Emanuel, R. (1976, January). Using LOGO to catalyse communication in an autistic child. *D.A.I. Resesarch Report No. 15*. University of Edinburgh.
- Watt, D. (1982). LOGO in the schools. *Byte*, 7, 116-134.
- Woodworth, R., & Schlosberg, H. (1954). *Experimental psychology*. New York: Holt, Rinehart & Winston.

## ESSAY REVIEWS

### II

#### SELF-FULFILLING PROPHECIES IN THE CLASSROOM: DEAD END OR PROMISING BEGINNING?

Valerie S. Smead

*Western Illinois University*

Teacher expectancy research has been grinding along since the initial burst of enthusiasm in the late sixties and early seventies. Laboratory studies artificially manipulating teachers' expectations still appear (Kedar-Voivodas & Tannenbaum, 1979; Taylor, 1979; Ysseldyke & Foster, 1978). Naturalistic studies continue (Good, Cooper, & Blakey, 1980; Luce & Hoge, 1978). Continued scrutiny also reveals further methodological and theoretical problems (Dusek, 1975; Reschly & Lamprecht, 1979; Rogosa, 1980; West & Anderson, 1976).

Originally, such research was in tune with the sociopolitical climate of the time. *Pygmalion in the Classroom* (Rosenthal & Jacobson, 1968) appeared during a period of liberalism and political unrest. Helping efforts at such times tend toward adjusting social institutions to meet individual needs (Levine & Levine, 1970). Broad change is seen as desirable and achievable. The notion that teacher expectations were responsible for students' differential achievement outcomes fit in well. If the teacher got what he or she expected, rather than what the students were capable of producing, social injustice was being perpetuated in part by the educational system. Teachers expecting certain children to do well (perhaps white or upper SES children) or badly (perhaps minority or lower SES children) were producing what they anticipated in the absence of real differences initially. Teachers and school systems, not children, were in error.

Now the temper of the time has changed. According to a recent *Reader's Digest* article, "Why Kids Can't Write," teachers are referring to the last decade as the "Bad Period" in education (Brandt, 1981). Considerably milder interpretations of

---

Dr. Smead is an Assistant Professor of Psychology at Western Illinois University, where she is also a staff member of the Psychology Clinic. Her teaching and research areas are in community mental health, applied psychology, and human development.



teacher expectations as sustaining rather than causing student achievement differences are also beginning to appear (Cooper, 1979; Good et al., 1980). While not minimizing the significance or need for social concern implied, Cooper (1979), in reviewing the literature, maintains:

For now, then, it seems best to conclude that expectations influence performance, but they likely sustain it at a preexisting level or allow innate differences in student performance to emerge rather than radically alter its course. (p. 392)

It will be argued here that despite methodological difficulties, competing rival hypotheses, and a different political climate, self-fulfilling prophecies in the classroom are still potentially important phenomenon deserving further study. Based on a review of selected research, an interactional framework for conceptualizing the effect will be presented. It does not cast teachers or the educational system as villains, nor students as passive victims. Rather, the interaction over time between teachers, students, and educational system is used as the focus for investigation.

Interactional approaches are not new in conceptualizing the teacher expectancy effect (see Braun, 1976; Cooper, 1979; Finn, 1972; Good, 1980) or in educational research in general (see Hobbs, 1975; Smead, 1977; Tracy, Gibbins, Kladder, & Daggy, 1974; Winne & Marx, 1977). Perhaps this particular effort will prove useful because it attempts to acknowledge and build upon the complex interactions which have been implicated in expectancy research. Ongoing reciprocal relationships between teacher and student are not seen as methodological dilemmas, but as the focus for investigation.

First, the research will be reviewed generally to provide an historical perspective which highlights some of the difficulties that researchers have confronted. Next, findings will be presented within an interactional framework. Both teacher and child effects on the relationship between expectations, behavior, and achievement will be considered. Other expectations in the classroom will also be addressed. This will be followed by speculations on stalking the expectancy effect over the course of a child's school career.

#### *Strengths and Weaknesses of the Current Data Base*

Though there have been instances where meager or negative results have been reported (Fleming & Anttonen, 1971; José & Cody, 1971; Humphreys & Stubbs, 1977; Luce & Hoge, 1978), many studies targeted towards various aspects of expectancy transmission have reported significant effects. Investigations utilizing videotapes (Foster, Ysseldyke, & Reese, 1975; Mason & Larimore, 1974; Ysseldyke & Foster, 1978), hypothetical instances (Bar-tal & Saxe, 1979), short term laboratory teaching situations (Beez, 1970), and "phantom" students (Taylor, 1979) appear to have a particularly good batting average. In addition, more naturalistic studies have reported correlational relationships between teacher expectations and student achievement or ability in actual classroom situations (Brophy & Good, 1970; Dusek & O'Connell, 1973; O'Connell, Dusek, & Wheeler, 1974; Palardy, 1969; Seaver, 1973; Willis, 1973). Such studies have also led to documentation of links between teacher expectations and teacher and student behavior (Brophy & Good, 1970; Good et al., 1980). In fact, when all the studies of expectation (teacher, experimenter, laboratory, field) are combined in a meta-analysis of the expectancy effect, "the reality of the phenomenon is beyond doubt and the mean size of the effect is clearly not trivial" (Rosenthal & Rubin, 1978, p. 385).

Unfortunately, there are problems with interpreting these data which could limit the seemingly impressive implications. Definite strengths in terms of the psycho-

logical credibility of the concept (Braun, 1976) and the amount of research it has stimulated are difficult to deny. Documentation that expectancies can be induced in the laboratory and that naturally occurring expectations relate to behavior and achievement in the field also may be beyond dispute. However, compelling demonstrations that teachers actually bias (see Dusek, 1975; West & Anderson, 1976), students' learning success, due to inaccurate expectations, are not as readily available.

The results of laboratory studies utilizing videotapes, hypothetical instances, phantom students, and even brief interactions are difficult to generalize to the ongoing interactions between teachers and students which occur in school situations. As Mason and Larimore (1974) warned more than a decade ago:

This study is a laboratory situation and teachers may not interact with TV sets and psychology reports in the same way that they interact, with children and psychological reports in real situations. (p. 48)

In some cases such studies are questionable even as laboratory demonstrations (Dusek, 1975; Reschly & Lamprecht, 1979). In the absence of information except expectancy manipulations and brief interaction with a videotape or a stranger, it is not particularly surprising that subjects accept and act on the information which they have received. Which piece they chose is of interest and usually intended to generate hypotheses to be tested against actual teachers in real classrooms. It is not a demonstration that: (1) teachers actually form expectations based on labels, psychological reports, cumulative folders, etc. which override student capacities as demonstrated in their classrooms; (2) such information is necessarily an invalid source for forming expectations; perhaps not all psychological reports and diagnostic labels represent "bogus" information; or (3) in those cases where the information is invalid, that teachers will create prophecy effects through their own credulity. One investigation which extended teacher exposure time to videotaped student expectee beyond the usual 10 to 12 minutes reported that, in fact, the effect dissipated after 30 minutes (Reschly & Lamprecht, 1979). Initially, teachers accepted the expectations they had been given. However, after watching the child for just one half-hour, the induced expectations lost their hold.

Studies in the field utilizing naturally occurring teacher expectations bypass some of the difficulties described above. They are based upon ample interaction and actual classroom behavior. However, correlational relationships between teacher expectation and student ability or achievement are subject to another difficulty in interpretation. As Crano & Mellon (1978) suggest, documenting such relationships may "limit the outcome of the research to, at best, a statement of the obvious" (p. 40). That is, competent teachers would be expected to accurately forecast student achievement. A significant relationship between expectation and achievement, therefore, may reflect an expectancy effect, or astute teachers, or some combination of the two (see Crano & Mellon, 1978; Dusek, 1975; West & Anderson, 1976).

Attempts to address this issue have yielded uneven results and further methodological difficulties. One approach involves artificial manipulation of teacher expectation, as did the original Rosenthal and Jacobson study. Such a move beyond correlation to experimental control, though appealing, again raises questions concerning generalizability. The demand characteristics of supposedly correct data, supplied by supposedly knowledgeable and honorable researchers might, on occasion, override teachers' better judgments for a time. Such situations do not typify day-to-day life in the school. In some cases, significant effects have occurred



through manipulation of expectation (Meichenbaum, Bowers, & Ross, 1969; Zanna, Sheras, Cooper, & Shaw, 1975). In others, teachers have indicated that they did not, in fact, find manipulated test results credible (Fleming & Anttonen, 1971; José & Cody, 1971). In one instance where naturally occurring teacher expectations were compared to an attempt at manipulation, only those occurring naturally related to student achievement (Dusek & O'Connell, 1973). This correlational relationship held on three different testing occasions during the school year (Dusek & O'Connell, 1973) and even after the children had gone on to new classrooms the next year (O'Connell, Dusek, & Wheeler, 1974). For further critiques of artificial manipulation see Dusek (1975), Finn (1972), Pelligrini and Hicks (1972).

Another tack involves detecting inappropriate expectations and relating them to subsequent achievement (Palardy, 1969; Sutherland & Goldschmid, 1974). However, with young students without a history of test scores, grades, etc., making an accurate determination of ability, against which expectations may be evaluated, is difficult. It is questionable whether a pre-reading screening test (Palardy, 1969) or even four WISC subtests and a group intelligence test (Sutherland & Goldschmid, 1974) given early in elementary school are sufficient bases for determining the accuracy of teacher expectations. The significant effects which emerged may document prophecy effects for inaccurate teacher expectations, or they may document that teachers were, in actuality, better predictors of future performance than the measuring devices utilized.

A final approach involves utilizing cross-lagged panel designs in an attempt to determine causal predominance. The question addressed is, are expectations more likely to be accurate reflections of previous achievement or may achievement more appropriately be considered a reflection of previous expectations? Proponents of this technique maintain that with two correlated variables such as expectation and achievement, this procedure allows for determination of the direction of causality. Thus far, in one instance teacher expectation was supported as a causal variable in terms of elementary students' achievement (Crano & Mellon, 1978); in another such support was not obtained in terms of high school students' grades (Humphreys & Stubbs, 1977). In addition, some maintain that cross-lagged panel designs are totally inadequate methods for attempting to unravel causality and should be abandoned (Rogosa, 1980). Utilizers of this technique do acknowledge that if a third variable were operating on both variables to unequal extents and went by undetected, predominance would inappropriately be assigned to one of the original variables (Crano & Mellon, 1978; Humphreys & Stubbs, 1977).

In sum, despite an impressive array of significant results, it is difficult to demonstrate conclusively that teacher expectations bias student learning. Expectations which have been artificially manipulated in the laboratory or in actual school situations are subject to questions concerning generalizability. At times, manipulations may not be believed by teachers, who may be operating on the basis of their own naturally occurring expectations. Credibility issues are becoming more important in the laboratory, as well. At the end of a recent experiment more than half of the subjects indicated the suspicion that it related in some way to teacher expectations (Taylor, 1979). Studies which forego experimental control through manipulation are confronted with the problem that naturally occurring expectations may be accurate. If such is the case then teacher expectations are not biasing student learning, but rather correctly forecasting it.



### *Expectations in an Interactional Framework*

Starting with the oft-documented correlation between teacher expectations and student achievement, a different question might be asked. Rather than further attempts to determine how much expectations influence (rather than are influenced by) student ability and achievement, a bi-directional relationship might simply be accepted. In addition to the appeals of common sense and compromise, until totally accurate measures of child academic potential are devised for preschool children, teacher expectation cannot be separated from student achievement and potential. That is, once exposed to teachers who behave in line with their expectations, or alternatively, expect in line with their behaviors (see Brophy & Good, 1970; Good et al., 1980; West & Anderson, 1976), the effects of such treatment on standardized test performance cannot be discounted. High teacher expectancy students may learn more. Unless ability is seen as a fixed attribute, they may even increase their potential for learning in contrast to lower expectancy students. Over the course of a child's school career, in some cases, expectation, behavior, and achievement, therefore, might be expected to follow upward or downward spirals. Once a child has entered kindergarten, teacher expectations and behaviors may be reflected in his or her learning success, grades, and standardized test performance which, in turn, may influence the next teacher's expectations and behaviors (see Finn, 1972, and Good, 1980, for somewhat similar discussions).

From this perspective, instead of further attempts to unravel these variables, attempts to chart, understand, and eventually change the course of their covariation during a child's school career are suggested. The question becomes why, for some children, do expectations and achievements begin a downward spiral and how may such a spiral be prevented.

In addressing this question, several additional areas of expectancy research must be considered. First, academic expectations and behavior in the classroom will be discussed. Student as well as teacher expectations and behaviors have surfaced as meaningful variables here. Secondly, teachers and students may have expectations which go beyond the assumed academic potential of the student. Finally, time and situation may influence the interaction between expectations, behavior, and achievement. Each will be discussed in turn.

### *Expectations and Behavior*

The link between expectations and classroom behavior has received intensive investigation (see Brophy & Good, 1970; Cooper, 1979; Good, 1980; Good et al., 1980). As mentioned, several teaching behaviors which covary with teachers' expectations have been uncovered. For instance, in some cases, teachers wait less time for low expectancy students' responses, demand less from them, criticize more, and seat them further from the teacher's desk. Good et al. (1980) report on these and seven other differential teaching behaviors which have been reported in at least two studies each. Such differences do not, however, occur for all teachers. Based on a large pool of observational studies done in one geographic region, Good (1980) estimates that approximately one-third of the teachers acted in ways which could have exaggerated differences between high and low students unnecessarily.

At this point students must also be considered. Do "good" and "bad" students treat teachers differently? Do children see themselves as good and bad students and act on their own expectations? The answer may be affirmative on both counts. More than a decade ago in a correlational study utilizing first graders, it was re-

ported that “the highs seek out the teacher and initiate interactions with her more frequently than the lows . . . the highs much more frequently show their work . . . or ask questions about it” (Brophy & Good, 1970, p. 368). This possibility of student as active agent in the expectancy drama was initially pursued by only a few (Zanna et al., 1975). Perhaps these few were impervious to the tenor of the times, or perhaps aware earlier than many that the position of passive victim is no more comfortable than that of villain. One early attempt at theory building included student expectation as a central variable (Finn, 1972). Several studies manipulated student as well as teacher expectation (Kern, 1973; Mulligan, 1973; Rappaport & Rappaport, 1975; Zanna et al., 1975). In the latter three studies, student expectation appeared to have either a similar or greater impact than that of the teacher.

More data have since accumulated. Students’ naturally occurring expectations for their achievement in eighth grade mathematics class correlated with later achievement across two dependent measures and times of the year (Smead & Chase, 1981). Learning disabled students, it now appears, are operating within what Finn (1972) suggested was a network of expectations. The students, their mothers, and their teachers all hold low expectations for their future achievement; students themselves may react to success and failure in ways which facilitate the occurrence of such expectations (Bryan & Pearl, 1979). Even among regular second and third graders, those whom the teacher rated as being in the bottom third of the class academically rated themselves significantly lower than did those in the top third (Stipek, 1981). These self-ratings also correlated with ratings they received from their peers. Similarly, eighth grade mathematics students, particularly high expectancy students, reported substantial agreement between their own expectations and those they felt that their parents and peers held for them (Smead & Chase, 1981).

Finally, additional research confirms that students for whom teachers hold high and low expectations behave differently in the classroom. It is introduced with the caution:

It must be kept in mind that the data to be reported are correlational. It is impossible to determine, therefore, if differences in the classroom interaction are caused by teacher or student behavior (or if influence is bidirectional). (Good et al., 1980, p. 380)

Thus, child as well as teacher effects may be found in the ongoing interaction between expectations, behavior, and achievement.

### *Beyond Academic Potential*

Cooper (1979) has argued that teachers’ differential behaviors toward high and low expectancy students might reflect beliefs about how controllable such children are. For instance, criticizing low expectancy students more and praising them less might serve to discourage low control exchanges between the teacher and these students (see Cooper, 1979). Other work suggests that not only may achievement expectations contain associated beliefs about controllability, but that teachers and perhaps students may have expectations which are not based only upon assumed academic potential of the student. These expectations may influence such potential, however.

For instance, two types of expectations were measured in the Crano and Mellon (1978) study. In addition to teachers’ expectations for students’ academic performance, expectations for social performance (in terms of disobedience, pleasantness to have in class, and attitude toward school work, including work habits) were obtained. They reported that the “most positive and consistent results” (p. 47) in



terms of expectations impact upon later academic achievement, were obtained from these social expectations and concluded:

It is obvious . . . that the effects of social evaluations and expectations, based as they are on aspects of a child totally incidental to intellectual promise, represent a potential educational inequity of even greater import than that emphasized by Rosenthal and Jacobson. (1968, p. 48)

Elsewhere, factor analytic study of teacher expectations for a variety of child characteristics (conduct, school performance, sociability) suggested a first factor that was fairly similar for all teachers (Garner & Bing, 1973). This factor included attitudes toward work, school performance and intellectual potential. Interestingly, the positive aspects of many other dimensions loaded on this factor, "suggesting a fairly widespread stereotype of the 'good' pupil, the 'average' pupil, and the 'poor' pupil" (p. 36). In this instance a consistent relationship between expectations and verbal interchange, the dependent measure utilized, was not reported. Elsewhere, teacher rating of student motivation correlated significantly with a variety of teacher behaviors while teacher ratings of ability and achievement generally did not (Luce & Hoge, 1978). In response to hypothetical instances (Kedar-Voivodas & Tannenbaum, 1979), teachers were more influenced by descriptions of classroom behaviors than by sex, label, and therapy condition, each of which had no significant effect. In addition, acting out behaviors had more of a detrimental effect on teacher expectation than did withdrawn behaviors of equal severity.

Students, themselves, have expressed the belief, again in response to hypothetical instances, that student effort is more critical for teacher judgment than student ability (Harari & Covington, 1981). This view held for primary and secondary students. In the primary grades, the reputation of being a hard worker was also a student value. Based on further exploration with semistructured interviews, students in grades 1, 2, and 4 appeared to see effort and outcome as synonymous. Those who tried hard were expected to do well; in fact, they could even improve their ability. In actual classroom situations, early elementary students justified their judgments of peers' ability most frequently by referring to their work habits (Stipek, 1981).

In sum, it appears that child motivation, manageability, and general willingness to play the good student role may also influence achievement expectations and/or behaviors in the classroom.

### *Expectation, Achievement, and Behavior Across Time and Situation*

In attempting to chart the covariation of expectations, behavior, and achievement over the course of a school career, not only must differential teacher expectation and behavior be taken into consideration, but children, themselves, cannot be ignored. Their behavior, expectations and potential for playing the good student role may also contribute.

As discussed, once children enter kindergarten, the effects of teacher expectations and behaviors may be reflected in childrens' learning, grades, and standardized test performance. At this point, it appears that the teacher might also influence student behaviors and expectations which, again, might be carried into subsequent classrooms. However, the teacher is no longer clearly the starting point. Child differences in behavior may also cause teacher differential behaviors and perceptions. Child expectations, motivation, and manageability also may not be entirely contingent on teacher treatment. Therefore, some mix of teacher and child effects



on achievement and potential may be carried into new classrooms and aid in perpetuating upward and downward spirals. The issue which must be addressed is who and what is more important at different points in the child's school career. Unfortunately, with interactional data such as these, considerable speculation is involved in attempting to chart expectations across the course of a school career.

Teacher expectations and behavior may be more important starting with the early grades of elementary school than they will be later. Several investigators have mentioned this as a potentially critical time (Cooper, 1979; Good, 1980; Sutherland & Goldschmid, 1974). For many elementary school children the teacher's credibility rating is high; she or he is the one who "knows" (Braun, 1976). Also, at this age children are more impressionable and anxious to please adults (Good, 1980). Several studies documenting relationships between teacher expectation and student achievement, in fact, have involved these years (Brophy & Good, 1970; Crano & Mellon, 1978; Dusek & O'Connell, 1973; O'Connell, Dusek, & Wheeler, 1974; Palardy, 1969; Seaver, 1973). Longitudinal data across the first four grades, an interesting byproduct of one of the cross-lagged panel designs (Crano & Mellon, 1978) also suggest that during this period expectations may tend toward increasing consistency between teachers. In particular, the following correlations were obtained for teacher expectations (academic and social) .289 between teacher, grade one, and teacher, grade two; .550 between two and three; and .562 between three and four. Finally, West and Anderson (1976) have suggested that teacher expectation would be more effective when previous interaction is minimal, the content area is new and not too similar to previously learned material, and teacher behavior is essential for gaining knowledge. For many children, early elementary school may meet these criteria.

Downward or upward spirals need not start early in school, however, and if they do the teacher need not be the only active participant. It has been repeatedly documented that in some classrooms, teachers do not appear as expectancy prone as in others (Brophy & Good, 1970; Good et al., 1980; Good, 1980). Therefore, children might escape teachers who manifest differential behaviors toward high and low students. In addition, they might escape relegation to high and low categories. In one instance involving older students (6th to 8th graders), teachers indicated more confidence in their estimates of students who scored toward the extreme ends of the continuum (Felson & Bohrnstedt, 1980). This relationship was noted particularly with positive extremes. Thus, even in the case of expectancy prone teachers, if student performance is not extreme, but more towards the middle of the continuum, then differential expectations might be precluded.

Even at this age, child, parent, and peer contributions cannot be discounted. In extreme cases, second and third grade students may already manifest differential expectations for themselves (Stipek, 1981). Perhaps students who adopt behaviors associated with motivation, manageability, and the good student role may prove less susceptible to negative spirals than peers of similar ability who do not. Expectancy resistant parents might be another lever—they might rebut low grades, transmit essential knowledge, or mold good student behaviors themselves when child or teacher is in danger of initiating such a spiral. Peers who model appropriate behavior and attitudes and provide feedback based on work habits may also contribute.

As time goes on, teacher effect may be further diluted as children's own expectations take on more importance at the later elementary or secondary level. First, classroom behavior might be expected to be "progressively determined by student's

past experiences and performances in similar situations” (Good, 1980, p. 117). Secondly, as discussed below, views of their own abilities may be differentiating and solidifying.

Kifer (Note 1 reported by Newman & Newman, 1981) asked youngsters to compare their school ability to that of other children in their grades. Responses of A and B students in grades two through eight were compared to those of D and E students. In the early grades estimates of ability were not related to grades; the trend was for all students to express high self-concepts of ability. But as grade level increased there was a steady downward trend in expressed self-concept of ability for D and E students. By grade eight they were expressing quite low appraisals of their ability while A and B students continued to express high appraisals. Elsewhere, the validity of students’ expectations (based on specific performance tasks) was found to increase across first, third, and fifth graders; even at the fifth grade level validity did not exceed 1.8 on a 3-point scale (Clifford, 1975). Overestimation tendencies were reported as was failure to process task feedback.

Thus, not only may behaviors and attitudes towards school crystallize with time, but children may enter new classrooms with progressively more clear-cut expectations for themselves. They may also be changing their earlier view that “ability as a concept is fluid, malleable, and highly dependent on effort” (Harari & Covington, 1981, p. 25). In that investigation, by the sixth grade students no longer expressed the idea that effort causes ability change. By eighth grade ability was seen as a more stable factor and as a “strong moderator of the effort-outcome covariation” (p. 25). The view that anyone, irrespective of ability, who studies hard will do well was no longer endorsed. By eleventh grade outcome was seen primarily as a function of ability.

Finally, on entering junior high school or the later elementary grades in some schools, students are exposed to not one but several teachers. Many times these teachers are exposed to not 30 but over one hundred students. Often teacher behavior is no longer essential to accumulating knowledge—other sources are homework, reading the text, or copying a friend’s notes. The content area may not be novel, and previously formed expectations for reading, mathematics, and even physical education may have direct bearing on current student performance in the classroom.

However, just as early elementary teachers’ expectations and behaviors are not seen as omnipotent, neither are those of secondary students. Not only are new peer groups who express different expectations and attitudes possible, but so are teachers or subjects which suddenly inspire them. Students who might be most flexible at this stage are those who have not formed stable images of themselves as learners due to inconsistent performance or attitudes toward school. Also of interest is that group in the broad middle range of achievement who can acquire sufficient skills to “take off” if appropriately motivated from within or without.

### *Conclusion*

The question is not whether expectations bias achievement. This review suggests that child and teacher expectations, behavior, and child achievement work together over the course of a child’s school career. The question is which variables are most accessible to modification at which point in time. At this point, much speculation is involved in charting what might be called the expectancy-behavior-achievement effect. Longitudinal studies monitoring these variables from



kindergarten through secondary school are needed. At present, the possibility of a gradually tightening web of variables is suggested for some children. Those who do not play the good student role and score towards the lower end of the continuum may be among those particularly vulnerable to negative spirals.

Also suggested are multiple intervention points from child expectation, to teacher preference for particular behaviors, to preparing expectancy resistant parents. In the years since *Pygmalion in the Classroom*, the magical connotation of teacher expectations has disappeared. Now, at times, all involved appear to be potential "victims" (see Ryan, 1971, 1976)—teachers, parents, peers, and students who come to believe they are dumb. All appear to have some potential power in halting downward spirals, as well. Further experimental manipulations are necessary. However, rather than continuing ill-fated attempts to untie variables which naturally work together, these must address the impact of particular teacher, parent, or student-centered interventions at particular times in the child's school career.

#### Notes

1. Kifer, E. The impact of schooling on perceptions of self. Paper presented at the self concept symposium, Boston, September, 1978.

#### References

- Bar-tal, D., & Saxe, L. (1979). Teachers' information processing: Effect of information about pupils on teachers' expectations and affect. *Psychological Reports*, 44, 599-602.
- Beez, W. V. (1970). *Influence of biased psychological reports on "teacher" behavior and pupil performance*. Unpublished doctoral dissertation, Indiana University.
- Brandt, A. (1981). Why kids can't write. *Reader's Digest*, 118, 99-102.
- Braun, C. (1976). Teacher expectation: Sociopsychological dynamics. *Review of Educational Research*, 46, 185-213.
- Brophy, J. E., & Good, T. C. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology*, 61, 365-374.
- Bryan, T., & Pearl, R. (1979). Self concepts and locus of control of learning disabled children. *Journal of Clinical Child Psychology*, 223-226.
- Clifford, M. M. (1975). Validity of expectation—A developmental function. *Alberta Journal of Educational Research*, 21, 11-17.
- Cooper, H. M. (1979). Pygmalion grows up: A model for teacher expectation communication and performance influence. *Review of Educational Research*, 49, 389-410.
- Crano, W. D., & Mellon, P. M. (1978). Causal influences of teachers' expectations on children's academic performance: A cross-lagged panel analysis. *Journal of Educational Psychology*, 70, 39-49.
- Dusek, J. B. (1975). Do teachers bias children's learning? *Review of Educational Research*, 45, 661-684.
- Dusek, J. B., & O'Connell, E. J. (1973). Teacher expectancy effects on the achievement test performance of elementary school children. *Journal of Educational Psychology*, 65, 371-377.
- Finn, J. D. (1972). Expectations and the educational environment. *Review of Educational Research*, 42, 387-410.
- Felson, R. B., & Bohrnstedt, G. W. (1980). Attributions of ability and motivation in a natural setting. *Journal of Personality and Social Psychology*, 39, 799-805.



- Fleming, E. S., & Anttonen, R. G. (1971). Teacher expectancy as related to the academic and personal growth of primary-age children. *Monographs of the Society of Research in Child Development*, 36, Serial No. 145.
- Foster, G. G., Ysseldyke, J. E., & Reese, J. H. (1975). I wouldn't have seen it if I hadn't believed it. *Exceptional Children*, 41, 469-473.
- Garner, J., & Bing, M. (1973). The elusiveness of Pygmalion and differences in teacher-pupil contacts. *Interchange*, 4, 34-42.
- Good, T. L. (1980). Classroom expectations: Teacher and pupil interactions. In J. H. McMillan (Ed.), *The social psychology of school learning*. New York: Academic Press.
- Good, T. L., Cooper, H. M., & Blakey, S. L. (1980). Classroom interaction as a function of teacher expectations, student sex, and time of year. *Journal of Educational Psychology*, 72, 378-385.
- Harari, O., & Covington, M. V. (1981). Reactions to achievement behavior from a teacher and student perspective: A developmental analysis. *American Educational Research Journal*, 18, 15-28.
- Hobbs, N. (1975). *The futures of children*. San Francisco: Jossey-Bass.
- Humphreys, L. G., & Stubbs, J. (1977). A longitudinal analysis of teacher expectation, student expectation, and student achievement. *Journal of Educational Measurement*, 14, 261-270.
- José, J., & Cody, J. (1971). Teacher-pupil interaction as it relates to attempted changes in teacher expectancy of academic ability and achievement. *American Educational Research Journal*, 8, 39-49.
- Kedar-Voivodas, G., & Tannenbaum, A. J. (1979). Teachers' attitudes toward young deviant children. *Journal of Educational Psychology*, 71, 800-808.
- Kern, J. C. (1973). Experimenter effects in pupil testing (Doctoral dissertation, The City University of New York, 1972). *Dissertation Abstracts International*, 33, 3919B.
- Levine, M., & Levine, A. (1970). *A social history of helping services*. New York: Appleton-Century-Crofts.
- Luce, S. R., & Hoge, R. D. (1978). Relations among teacher rankings, pupil-teacher interactions, and academic achievement: A test of the teacher expectancy hypothesis. *American Educational Research Journal*, 15, 489-500.
- Mason, E. J., & Larimore, D. L. (1974). Effects of biased psychological reports on two types of teachers' ratings. *Journal of School Psychology*, 12, 46-50.
- Meichenbaum, D. H., Bowers, K. W., & Ross, R. R. (1969). *Journal of Personality and Social Psychology*, 13, 306-316.
- Mulligan, J. P. (1973). Teacher and student expectations and student achievement in college reading improvement courses (Doctoral dissertation, Rutgers University, 1972). *Dissertation Abstracts International*, 33, 3776A.
- Newman, P. R., & Newman, B. M. (1981). *Living: The process of adjustment*. Homewood: Dorsey.
- O'Connell, E. J., Dusek, J. B., & Wheeler, R. J. (1974). A follow-up study of teacher expectancy effects. *Journal of Educational Psychology*, 66, 325-328.
- Palardy, J. M. (1969). What teachers believe—What children achieve. *Elementary School Journal*, 69, 370-374.
- Pellegrini, R. J., & Hicks, R. A. (1972). Prophecy effects and tutorial instruction for the disadvantaged child. *American Educational Research Journal*, 9, 413-418.
- Rappaport, M. M., & Rappaport, H. (1975). The other half of the expectancy equation: Pygmalion. *Journal of Educational Psychology*, 67, 531-536.
- Reschly, D. J., & Lamprecht, M. J. (1979). Expectancy effect of labels: Fact or artifact? *Exceptional Children*, 45, 55-58.
- Rogosa, D. A. (1980). A critique of cross-lagged correlation. *Psychology Bulletin*, 88, 245-258.

- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston.
- Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *The Behavioral and Brain Sciences*, 3, 377-415.
- Ryan, W. (1971, 1976). *Blaming the victim*. New York: Vintage Books.
- Seaver, W. B. (1973). Effects of naturally induced teacher expectancies. *Journal of Personality and Social Psychology*, 28, 333-342.
- Smead, V. S. (1977). Ability training and task analysis in diagnostic/prescriptive teaching. *The Journal of Special Education*, 11, 113-125.
- Smead, V. S., & Chase, C. I. (1981). Student expectations as they relate to achievement in eighth grade mathematics. *Journal of Educational Research*, 2, 115-120.
- Stipek, D. J. (1981). Children's perceptions of their own and their classmates' ability. *Journal of Educational Psychology*, 73, 404-410.
- Sutherland, A., & Goldschmid, M. L. (1974). Negative teacher expectation and IQ change in children with superior intellectual potential. *Child Development*, 45, 852-856.
- Taylor, M. C. (1979). Race, sex, and the expression of self-fulfilling prophecies in a laboratory teaching situation. *Journal of Personality and Social Psychology*, 37, 897-912.
- Tracy, M. L., Gibbins, S., Kladder, F. W., & Daggy, N. (1975). *Case conference: A simulation and source book*. Ann Arbor: Distributed by the Publication and Distribution Service of the University of Michigan Press.
- West, C. K., & Anderson, T. H. (1976). The question of preponderant causation in teacher expectancy research. *Review of Educational Research*, 46, 613-630.
- Willis, S. L. (1973). Formation of teachers' expectations of student academic performance (Doctoral dissertation, University of Texas, Austin, Texas, 1972). *Dissertation Abstracts International*, 33, 4960A.
- Winne, P. H., & Marx, R. W. (1977). Reconceptualizing research on teaching. *Journal of Educational Psychology*, 69, 668-678.
- Ysseldyke, J. E., & Foster, G. G. (1978). Bias in teachers' observations of emotionally disturbed and learning disabled children. *Exceptional Children*, 44, 613-615.
- Zanna, M. P., Sheras, P. L., Cooper, J., & Shaw, C. (1975). Pygmalion and Galatea: The interactive effect of teacher and student expectancies. *Journal of Experimental Social Psychology*, 2, 279-287.

## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by authors' last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary* or to the *Dictionary of Canadian English*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.





# **aer**

THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH

VOLUME XXX

NUMBER 3

SEPTEMBER 1984

PUBLISHED BY  
THE UNIVERSITY OF ALBERTA • EDMONTON

# THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

Published in March, June, September, December by the  
Faculty of Education, The University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

EDITOR: C. S. Bumbarger

## CONSULTING EDITORS

D. C. Bjorkquist  
University of Minnesota

P. E. Bryant  
University of Oxford

B. M. Chung  
Chungbuk University,  
Republic of Korea

A. Côté  
University of Ottawa

D. Fitzgerald  
The University of New England

D. B. Harrison  
The University of Calgary

N. L. Hersom  
The University of Saskatchewan

M. Laferrière  
McGill University

T. Morrison  
University of Manitoba

J. Richardson  
University of New Brunswick

P. Robinson  
Atlantic Institute of Education

T. Rogers  
University of British Columbia

R. E. Traub  
Ontario Institute for  
Studies in Education

K. Watson  
University of Reading

R. J. Wilson  
Queen's University

*The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, *Social Science Citation Index*, *ERIC/Current Index to Journals in Education*, and *Alberta Education Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, *Sociology of Education Abstracts*, *Language Behavior Abstracts*, *Cognitive Development Abstracts*, *Research into Higher Education Abstracts*, *School Organization and Management Abstracts*, and *Multicultural Education Abstracts*.

The subscription rate is \$12.00 per year; single copies are \$3.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, The University of Alberta, Edmonton, Canada, T6G 2G5.





ISSN 0002-4805

# The Alberta Journal of Educational Research

Volume XXX, Number 3

September, 1984

## CONTENTS

- Factors Contributing to the Job Satisfaction of School Principals ..... 157  
*D. Friesen, E. A. Holdaway and A. W. Rice*
- The Achievement Productivity of Psychological Variables:  
A Replication and Extension in Rajasthan, India ..... 171  
*H. J. Walberg, R. P. Singh and S. L. Tsai*
- Integrating School Curricula with Social Concerns ..... 179  
*F. J. Maynes and J. A. Ross*
- The Influence of Improved Reading Skill On Writing Skill ..... 194  
*J. Belanger and R. G. Martin*
- The Relationship Between Academic Achievement, Self-Concept,  
Creativity, and Teacher Expectations Among Native Children  
in a Northern Manitoba School ..... 213  
*W. E. Rampaul, M. Singh and J. Didyk*
- Prospective Teachers: Views of Programs and Needs ..... 226  
*M. T. Nixon and C. S. Bumbarger*

FACULTY OF EDUCATION  
*The University of Alberta*

## ANNOUNCEMENT

As of September 1, 1984, the Editor of *The Alberta Journal of Educational Research* will be Dr. Walter H. Worth, replacing interim editor, Dr. C. S. Bumbarger.

Correspondence regarding manuscripts should be directed to:

Dr. W. H. Worth, Editor  
The Alberta Journal of Educational Research  
Publication Services  
4-116 Education North  
University of Alberta  
Edmonton, Alberta  
T6G 2G5

---

## CALL FOR NOMINATIONS

The National Council on Measurement in Education announces an award for an outstanding *technical* or *scientific* contribution to the field of educational measurement occurring during the years 1982, 1983 or 1984. The deadline for submission of nominations is February 1, 1985. One may either nominate his/her technical contribution or, with permission, someone else's. Those responsible for the technical contribution need not be members of NCME. To be considered for the award, four copies of a 3-5 page statement describing the technology, application area, and products or results of the effort should be sent to: Awards Committee Chair, c/o Anthony J. Nitko, Department of Educational Research Methodology, University of Pittsburgh, Pittsburgh, PA 15260.

Examples of technical contributions include, but are not limited to, innovative ways of solving practical or theoretical measurement problems, inventive instrument development techniques, creative testing procedures or products, and scientific contributions to measurement research methodology. Selection criteria are quality, innovativeness, and positive impact of the technology on the field of educational measurement.

The award will be presented at the NCME's 1985 annual meeting in Chicago.

DAVID FRIESEN

EDWARD A. HOLDAWAY

*The University of Alberta*

and

ALAN W. RICE

*Department of Education, New South Wales, Australia*

## Factors Contributing to the Job Satisfaction of School Principals

*Job satisfaction studies have focused primarily on production workers rather than on administrators. This article examines the job satisfaction of 410 school principals in Alberta. Overall and facet satisfaction of these principals are examined and related to current theories and previous findings. Eight satisfaction factors were identified through factor analysis. Regression analysis showed that the three factors contributing most to the overall satisfaction of the school principals were Responsibility and Autonomy, Principal-Teacher Work Involvement, and Liaison at District Level. Rapport with Students, Resource Adequacy, and Salary and Benefits did not contribute significantly to the overall job satisfaction of the principals.*

Job satisfaction has been studied in different kinds of organizations using a number of different approaches. According to Locke (1976, p. 1297), 3,350 articles or dissertations appeared on the subject from 1958 to 1976. The common objective of these studies has been to identify those facets in the work situation which are related to satisfaction on the assumption that increased satisfaction leads to better performance on the job. An extensive body of literature since 1976 indicates continuing interest in this area.

---

Dr. Friesen is Professor of Educational Administration at the University of Alberta. He has conducted research in such areas as occupational stress, job satisfaction, student subcultures and comparative educational administration.

Dr. Holdaway is Professor of Educational Administration and Director of the Office of Institutional Research and Planning at the University of Alberta. His research interests include satisfaction of educators, the organization of educational research, the research-policy relationship, and organizational analysis.

Dr. Rice has recently been appointed Deputy Principal, Prairiewood High School, N.S.W. He has had extensive experience in the administration of in-service education and curriculum services with the Department of Education, New South Wales, Australia.



Most of these studies have explored the satisfaction which production workers experience in their jobs, and have sought to identify the facets which are related to their overall satisfaction. Until recently, only a few studies have examined the facets that are related to job satisfaction of administrators. The importance of job satisfaction to administrators, as compared with production workers, can be viewed from two perspectives. One is related to the administrators' role of increasing job satisfaction for their subordinates. The other, which has received much less attention, has to do with the ways that administrators achieve satisfaction for themselves in the work situation. The study on which this paper is based sought information on this second theme with the problem being: What aspects of the work situation appear to contribute substantially to the job satisfaction of school administrators?

### *Theoretical Considerations*

The traditional interpretation of job satisfaction is that it arises out of all the feelings that individuals have about their work (Hoy & Miskel, 1982, p. 339). Where the sum of the satisfaction of such things as pay, the job itself, and advancement opportunities gives rise to feelings of satisfaction, the individual is said to be satisfied with the job. Where this total leads to dissatisfaction, the individual is seen as dissatisfied with the job. Expectancy theory extended this traditional interpretation by pointing out that an individual's expectations of the job are also important determinants of job satisfaction (Vroom, 1964; Miner, 1980, pp. 135-137). For example, when an individual has a high expectation for advancement, failure of a job to meet this expectation will lead to dissatisfaction.

Gruneberg (1976, p. xi) argued that "those proposing an expectancy theory usually regard overall job satisfaction as a function of satisfaction with the various elements of the job." He further stated that the Hygiene/Motivation theory developed by Herzberg (Herzberg, Mausner, and Snyderman, 1959) views factors causing satisfaction and dissatisfaction as arising from separate and distinct causes. The factors that contribute to job satisfaction are seen by Gruneberg to "include the intrinsic nature of the job and achievement" (p. xi).

A great deal of controversy still surrounds the appropriateness and applicability of Herzberg's Hygiene/Motivation Theory, much of which centers on Herzberg's choice of method in identifying peak experiences related to satisfaction or dissatisfaction, and on confusion resulting from his use of the term "motivators" as being synonymous with "satisfiers." Gruneberg, for example, made this observation:

Herzberg's work has come in for severe criticism recently from a number of sources. In particular a number of studies appear to show that the same factors can cause both satisfaction and dissatisfaction. Others, such as Schneider and Locke (1971), are critical of the methodology employed by Herzberg. Nevertheless there is little disagreement among theorists on the importance of the job itself as a major factor in job satisfaction for most individuals, and to this extent at least Herzberg's influence has been immense. (Gruneberg, 1976, p. xi)

Miner (1980, p. 102) reviewed the assessments of the Motivation/Hygiene Theory and concluded that it has had a "tremendous appeal . . . for practicing managers over the past 15 or 20 years" but that it has failed to "receive the support needed to confirm it, in spite of an extended period of testing and a great deal of research" (p. 103). He admitted that a "new thrust in research might . . . yield different conclusions" (p. 103). In this study, no attempt was made to resolve these difficulties.

### *Relationship Between Overall Satisfaction and Facet Satisfaction*

Theorists such as Lawler (1973) and Locke (1969) have conceptualized a relationship between overall job satisfaction and satisfaction with specific facets of the job. Locke (1969, p. 330) described overall job satisfaction as “the sum of the evaluations of the discriminable elements of which the job is composed.” Each facet or aspect of the job contributes differentially to overall job satisfaction. According to both Locke and Lawler, those aspects perceived by individuals to be more important to them contribute more to their overall job satisfaction than those aspects perceived by them to be less important.

In addition to this conceptual debate, methodological considerations have arisen in measuring overall job satisfaction. These concern the issue of using a summed measure of overall satisfaction versus a single measure. Locke (1969, p. 331) suggested that one approach is to sum the measures of satisfaction with specific job aspects. Wanous and Lawler (1972) investigated nine different techniques used for combining measures to assess overall job satisfaction. The validity of calculating overall job satisfaction by such combining was, however, questioned by Van Maanen and Katz (1976). On the basis of their findings, they contended that there may be occasions when overall job satisfaction would not be consistent with a summed measure of satisfaction with specific facets of the job. Smith, Kendall, and Hulin (1969) also argued that a separate measure ought to be used when the contribution of each of the facets of the job to overall job satisfaction is to be assessed. Another advantage for a separate measure, as Smith et al. (1969) noted, is that it transfers the problem of weighting the various job facets according to their importance directly from the researcher to the respondent. In the study reported here, overall job satisfaction was measured independently of satisfaction with job facets.

### *Relationship Between Specific Work Variables and Job Satisfaction*

The relationship between specific work variables and job satisfaction has been addressed by several researchers. Their opinions are presented here because of their relevance to this study. For example, Vroom (1964, pp. 172-174) identified seven job facets as being common to different studies on job satisfaction:

1. attitudes toward the company and company management;
2. attitudes toward promotional opportunities;
3. attitudes toward job content;
4. attitudes toward supervision;
5. attitudes toward financial rewards;
6. attitudes toward working conditions; and
7. attitudes toward co-workers.

Whereas Vroom's list relates to both Herzberg's satisfiers and dissatisfiers, other writers emphasize the variables which are associated most strongly with satisfaction. For example, Lawler stated that the following job characteristics positively influence an individual's job satisfaction attitudes:

[the job] (1) allows him to feel personally responsible for a meaningful portion of the work, (2) provides outcomes that are intrinsically meaningful or are otherwise experienced as worthwhile, and (3) provides feedback about what is accomplished. (Lawler, 1973, p. 158)



Herzberg et al. (1959) and Porter, Lawler, and Hackman (1975) claimed that intrinsic aspects of the work are positively related to overall job satisfaction, provided that the workers in question have a strong desire to satisfy Maslow's higher order needs.

An additional variable, interaction with others both inside and outside the organization on work-related matters, was identified by Iannone (1973), Schmidt (1976), and Van Maanen and Katz (1976) as a vital component of the activities of managers and school principals. Further, Vroom (1964) noted that specific factors frequently emerge in which interaction is the key component. He suggested that relationships tend to be satisfying to the extent that there is congruence of attitudes, acceptance by others, and progress towards the attainment of goals.

Concerning extrinsic aspects, Van Maanen and Katz (1976) noted that labor relations specialists and union leaders have emphasized the importance of organizational policies and conditions of work as influences on overall job satisfaction. Lawler (1973) stated that the relationship of extrinsic reward to overall job satisfaction may vary according to the importance that the individual places on these rewards. Schmidt (1976), using a modified "Herzberg critical incident technique" with 74 secondary school principals and other administrators in Chicago, found that recognition, achievement, and advancement were perceived to be major determinants of these principals' overall satisfaction, whereas salary, interpersonal relationships, policy and administration, and supervision tended to be highly dissatisfying.

Recent studies with teachers in Alberta by Holdaway (1978) and superintendents in California by Cochran (1977) have identified job factors related to the satisfaction of educators. Cochran found that personal satisfaction with the superintendency had the highest mean score among the seven factors which he identified. He also found that job satisfaction of the superintendents was unrelated to whether their assignment was in an elementary, secondary or unified school district. The results of these and other studies led to identification of the following seven factors as being relevant to the job satisfaction of principals:

1. the work itself;
2. occupational status and prestige;
3. interaction with district administration;
4. interaction with teachers;
5. interaction with students;
6. salary and benefits; and
7. working conditions.

### *Research Questions and Methodology*

The studies discussed above guided the development of the questionnaire and the research methodology used in the study reported in this article. The following major research questions were addressed:

1. What is the extent of the overall satisfaction experienced by high school principals?
2. What clustering of items is produced when the principals' responses are subjected to factor analysis?
3. To what extent are the factor scores correlated with the overall satisfaction scores?



4. What relationships exist between selected background variables and the factor scores identified in (2) above, and with the overall job satisfaction scores?
5. What do the findings of this study contribute to the theory on job satisfaction?

In this study “facets” refers to aspects that have been identified in the literature as related to job satisfaction, as well as to experiences identified by respondents as leading to job satisfaction or dissatisfaction.

### *Sample*

All principals in the Province of Alberta (excluding principals of private, special and one-room schools) served as the population of the study. A stratified random sample was drawn to ensure proportional representation of principals from the four types of publicly funded administrative units. One-third of the names were selected producing a sample of 410 principals. Twenty-three returns of the 350 received were deleted because they were incomplete, leaving 327 (79.8%) usable responses.

### *Questionnaire*

Section A contained questions regarding various characteristics of the principals and their work setting including age, sex, marital status, and employment status of the spouse. Other questions sought information of the professional characteristics of the respondents, such as the level of education. Another set of questions sought information about the organizational setting in which the respondents worked, for example, the size of the school.

Overall job satisfaction was initially assessed in Section B by four items adapted from the studies of Thronson (1969), Johnson and Weiss (1971), and Holdaway (1978). Respondents were asked to rate the following four items on a six-point scale from “Highly Satisfied” to “Highly Dissatisfied”:

1. the effectiveness of your school in educating students who come to it (compared with other schools known to you);
2. social relationships in your work;
3. the chance to do something that makes use of your abilities; and
4. your overall satisfaction with your job.

Originally the measure of overall job satisfaction was to be the mean score of the four items in the Overall Satisfaction Instrument, but the Pearson correlation coefficient between Satisfaction with School Effectiveness and each of the other three variables was low ( $r=.20$  to  $.26$ ). Consequently, Item 4, Your Overall Satisfaction with Your Job, was selected for further analysis. The relationship between this variable and the mean score of the other three items in the Overall Satisfaction Instrument was both statistically significant ( $p<0.01$ ) and substantial ( $r=0.84$ ). Thus item 4 was used as the measure for Overall Job Satisfaction.

The review of the literature identifies seventy facets as being related to various aspects of the principal’s job. Several of these facets were adapted from the Minnesota Satisfaction Questionnaire reported by Johnson and Weiss (1971), while others were adapted from the questionnaire developed by Holdaway (1978). These seventy facets were rated on a six-point scale by a class of Master’s students in educational administration and were also examined by a panel of three doctoral candidates. Questions involving facets were deleted if the facets were judged to be of minor importance to the principal, if they were ambiguous and difficult to answer, or if they failed to discriminate among respondents. The remaining forty-five facets

were organized into the five categories of Working Conditions, Personnel-related Matters, School-related Matters, District-related Matters and Occupation-related Matters to constitute Section C of the questionnaire.

A six-point scale, which ranged from "Highly Satisfied" to "Highly Dissatisfied," was used to rate the extent of facet satisfaction. No provision was made for Undecided or Neutral or Not Applicable response categories. This approach was in accord with the view of Porter et al. (1975, p. 53) who stated that "people are rarely neutral about things they perceive or experience" and that they "tend to evaluate most things in terms of whether they like or dislike them."

Two sections of the questionnaire, Section B (Overall Satisfaction) and Section C (Facet Satisfaction) were pilot-tested by using the responses of twenty-four school principals. One item was subsequently deleted, one was added, and some wording modifications were made. A measure of internal consistency was obtained for the 45-item scale on facet satisfaction with Kuder-Richardson formula 20 (K-20) which revealed a reliability coefficient of 0.94.

### *Findings*

#### *Question 1: Overall and Facet Satisfaction of Principals*

About 28 percent of the principals indicated that they were highly satisfied with their work, 57 percent were moderately satisfied, and 10 percent slightly satisfied. About 5 percent of the principals indicated that they were at least slightly dissatisfied with their position. Nonetheless, 95 percent of the principals expressed satisfaction, the mean overall satisfaction score being 5.05 on the six-point scale.

This high level of job satisfaction was also apparent in the responses to the individual items in Section C of the questionnaire as shown in Table 1. The six items with the highest ratings of "satisfied" were relationships with teachers, relationships with students, social position in the community, consequences of participative staff decisions, competence of staff in coping with day-to-day problems, and sense of accomplishment as an administrator.

The least mean satisfaction was experienced with the way in which consultation between board and teachers concerning working conditions was conducted, the portion of time devoted to operational duties, the availability of counselling services, and the way in which collective bargaining was conducted.

#### *Question 2: Results of Factor Analysis*

The 45 job-satisfaction items of the questionnaire were factor-analyzed using the method of principal factoring with iterations (PA2) from the Statistical Package for the Social Sciences (SPSS), yielding eight orthogonal factors from a principal-axis varimax solution. Table 2 presents the factor loadings of the 45 items on the eight factors, accounting for 56.5 percent of the total variance. Inspection of the items with the highest loadings on the factors suggested names for the factors. The factors in order of percentage of total variance were named Liaison at District Level, Principal-Teacher Work Involvement, Status Recognition, Responsibility and Autonomy, Resource Adequacy, Salary and Benefits, Task Demands, and Rapport with Students.



TABLE 1  
ITEMS FOR WHICH THE HIGHEST AND LOWEST PERCENTAGES  
“SATISFIED” WERE OBTAINED  
(N=327)

| Item  | Percentage<br>Satisfied | Mean |
|---|-------------------------|------|
| Your relationships with teachers  | 97.9                    | 5.24 |
| Your relationships with students  | 96.6                    | 5.21 |
| Your social position in the community   | 94.2                    | 4.91 |
| Consequences of participative staff decisions   | 93.3                    | 4.86 |
| Competence of your staff in coping with<br>day-to-day problems  | 93.0                    | 4.98 |
| Your sense of accomplishment as an<br>administrator   | 92.7                    | 4.77 |
| The way in which teacher/board collective<br>bargaining is conducted                                      | 59.0                    | 3.57 |
| Availability of counselling services  | 56.0                    | 3.62 |
| The portion of time devoted to operational<br>duties  | 54.7                    | 3.62 |
| The way in which consultation between board<br>and teachers concerning working conditions<br>is conducted | 44.0                    | 3.22 |

*Question 3: Correlation of Factor Scores and Overall Satisfaction Scores*

The relationship of the factor scores to Overall Job Satisfaction was further examined through multiple regression analysis, controlling for the effect of background variables which previous researchers had considered relevant to the studies on satisfaction. These background variables were sex, age, educational level and experience of principals, and size and location (urban, town, rural) of the schools. Normally, the variable that contributes most to the variance in the criterion variable enters first. However, by specifying that all background variables be entered first, regardless of their contribution to Overall Satisfaction, the variance attributed to these variables was controlled. The factor scores were then entered in order of their contribution to Overall Satisfaction. Results of the regression analysis are summarized in Table 3. The selected background variables accounted for 4 percent of the variance in Overall Satisfaction. The discussion of this relationship is contained in the analysis of Question 4. Five factors accounted for most of the additional variance in Overall Satisfaction: Responsibility and Autonomy at 14 percent, Principal-Teacher Work Involvement at 8 percent, Liaison at District Level at 5 percent, Status Recognition at 3 percent, and Task Demands at 2 percent. Three factors—Salary and Benefits, Resource Adequacy, and Rapport with Students—failed to account significantly for any of the variance in Overall Satisfaction of Principals. The results indicated clearly that three factors accounted for most of the variance in Overall Satisfaction when the analysis was controlled for the effect of selected background variables. Those three factors were Responsibility and Autonomy, Principal-Teacher Work Involvement, and Liaison at District Level.



Question 4: Relationship of Satisfaction Scores to Background Variables

The correlation matrix between the satisfaction scores and the background variables is presented in Table 4. Analysis of the statistically significant correlations provides further elaboration of differences in satisfaction scores, especially as they relate to different principal characteristics and system characteristics. The analysis is summarized for each of the six background variables used in the study. In general, the correlation coefficients are very low.

*Educational level of the principals.* Only one correlation coefficient which was statistically significant ( $p<0.05$ ) was obtained between the educational level of the principals and their satisfaction factor scores. The higher the educational level of the principals, the more satisfied they tended to be with the factor of Resource Adequacy.

TABLE 2  
VARIMAX FACTOR SOLUTION FOR 45 SATISFACTION VARIABLES  
USING EIGHT FACTORS (N=327)

| Satisfaction Item  | Factors and Factor Loadings (: 100) |    |     |     |     |     |     |     | Communality (+ 100) |
|--|-------------------------------------|----|-----|-----|-----|-----|-----|-----|---------------------|
|  | 1                                   | 2  | 3   | 4   | 5   | 6   | 7   | 8   |                     |
| <u>Liaison at District Level</u>                         |                                     |    |     |     |     |     |     |     |                     |
| Method to implement policy                               | 80                                  | 02 | 13  | 09  | 05  | 03  | 09  | 01  | 68                  |
| Involvement in decision making                           | 76                                  | 10 | 12  | 19  | 05  | -05 | 09  | 04  | 66                  |
| Board expectations for principal                         | 74                                  | 01 | 15  | 20  | 03  | 04  | 16  | 15  | 64                  |
| Consultations on working conditions                      | 68                                  | 08 | 08  | -13 | 27  | 19  | -06 | -08 | 61                  |
| Relationships with central office                        | 58                                  | 04 | 09  | 27  | 06  | -09 | 10  | 12  | 46                  |
| Evaluation of principals                                 | 54                                  | 01 | 19  | 27  | -07 | 16  | 12  | 12  | 46                  |
| Availability of advice                                   | 52                                  | 10 | 20  | 27  | 13  | 00  | 08  | 17  | 44                  |
| Procedure on suspension and expulsion                    | 45                                  | 08 | 08  | -01 | 11  | 18  | 13  | 16  | 30                  |
| Authority of admin. position                             | 43                                  | 16 | 30  | 29  | 02  | 10  | 24  | 15  | 47                  |
| Collective bargaining procedures                         | 41                                  | 20 | -02 | -05 | 23  | 33  | -01 | 01  | 37                  |
| Opportunities for in-service education                   | 39                                  | 06 | 23  | 25  | 22  | 11  | 11  | 10  | 35                  |
| <u>Principal-Teacher Work Involvement</u>                |                                     |    |     |     |     |     |     |     |                     |
| Staff competence   | 00                                  | 69 | 13  | 02  | 14  | 08  | -01 | 11  | 53                  |
| Teacher willingness to adopt principal innovations       | -01                                 | 69 | 09  | 14  | -05 | 17  | 08  | 08  | 55                  |
| Relationships with teachers                              | 03                                  | 66 | 14  | 09  | 01  | 09  | 02  | 24  | 53                  |
| Staff attitudes to curriculum change                     | 21                                  | 64 | 08  | 11  | 19  | 07  | 09  | -05 | 52                  |
| Consequences of participative decisions                  | 12                                  | 60 | 24  | 31  | 03  | 10  | 08  | 15  | 58                  |
| Opportunity to help teachers solve professional problems | 09                                  | 41 | 12  | 17  | 09  | 19  | 28  | 17  | 37                  |
| <u>Status Recognition</u>                                |                                     |    |     |     |     |     |     |     |                     |
| Recognition by other professionals                       | 32                                  | 22 | 65  | 00  | 14  | 09  | 05  | 00  | 61                  |
| Recognition by others                                    | 21                                  | 15 | 59  | 24  | 00  | 27  | 16  | 10  | 58                  |
| Social position in community                             | 25                                  | 11 | 58  | 20  | 03  | 25  | 00  | 12  | 54                  |
| Parent attitude to education                             | 08                                  | 16 | 47  | -04 | 23  | 19  | 04  | 28  | 42                  |
| Sense of accomplishment                                  | 20                                  | 25 | 39  | 37  | 02  | 11  | 26  | 26  | 54                  |

TABLE 2 CONTINUED

| Satisfaction Item                              | Factors and Factor Loadings (: 100) |      |      |      |      |      |      |      | Communality (+100) |      |
|--|-------------------------------------|------|------|------|------|------|------|------|--------------------|------|
|  | 1                                   | 2    | 3    | 4    | 5    | 6    | 7    | 8    |                    |      |
| <u>Responsibility and Autonomy</u>             |                                     |      |      |      |      |      |      |      |                    |      |
| Freedom to allocate teaching assignments       | 33                                  | 32   | -01  | 56   | 20   | 04   | 15   | 8    |                    | 60   |
| Freedom to seek out new ideas                  | 22                                  | 23   | 12   | 52   | 20   | 09   | 01   | 25   |                    | 50   |
| Responsibility of the principal                | 29                                  | 24   | 26   | 48   | 07   | 09   | 26   | 15   |                    | 53   |
| Accountability for school program              | 39                                  | 30   | 12   | 47   | 16   | 21   | 07   | 06   |                    | 54   |
| Authority over budget preparation              | 28                                  | 12   | 12   | 42   | 26   | 11   | -01  | -04  |                    | 36   |
| Freedom to organize for individual differences | 16                                  | 28   | 07   | 37   | 31   | 07   | 16   | 24   |                    | 42   |
| <u>Resource Adequacy</u>                       |                                     |      |      |      |      |      |      |      |                    |      |
| Access to library resources                    | -01                                 | -03  | 08   | 15   | 61   | 18   | 10   | 14   |                    | 46   |
| Physical working conditions                    | 08                                  | 08   | 10   | 17   | 52   | 11   | 29   | -04  |                    | 42   |
| Availability of community facilities           | 14                                  | 11   | 28   | 03   | 51   | 18   | -01  | 07   |                    | 41   |
| Availability of clerical help                  | 34                                  | 06   | -06  | 15   | 41   | 03   | 30   | -11  |                    | 42   |
| Availability of counseling services            | 16                                  | 20   | 02   | 07   | 40   | 13   | 09   | 08   |                    | 26   |
| <u>Salary and Benefits</u>                     |                                     |      |      |      |      |      |      |      |                    |      |
| Retirement benefits                            | 01                                  | 09   | 15   | 04   | 03   | 57   | 11   | 01   |                    | 36   |
| Salary   | 15                                  | 15   | 16   | 01   | 13   | 54   | 10   | 02   |                    | 40   |
| Sabbatical leave                               | 01                                  | 14   | 07   | 05   | 18   | 54   | 12   | 05   |                    | 54   |
| Sick leave                                     | 07                                  | 06   | 15   | 19   | 19   | 50   | -03  | 16   |                    | 50   |
| <u>Task Demands</u>                            |                                     |      |      |      |      |      |      |      |                    |      |
| Portion of time devoted to operational duties  | 29                                  | 05   | 01   | 00   | 39   | 23   | 58   | 09   |                    | 64   |
| Number of hours expected to work               | 25                                  | 16   | 10   | 13   | 21   | 34   | 56   | 06   |                    | 59   |
| Variety of tasks in workload                   | 29                                  | 12   | 38   | 19   | 21   | 00   | 44   | 06   |                    | 52   |
| Effect of job on personal life                 | 17                                  | 25   | 31   | 10   | 21   | 17   | 41   | 14   |                    | 46   |
| <u>Rapport with Students</u>                   |                                     |      |      |      |      |      |      |      |                    |      |
| Relationships with students                    | 11                                  | 17   | 14   | 16   | 03   | 09   | 05   | 62   |                    | 48   |
| Attitude of students toward staff              | 16                                  | 33   | 13   | 11   | 12   | 05   | 09   | 59   |                    | 53   |
| Provision of custodial services                | 06                                  | 28   | -09  | 15   | 22   | 22   | 09   | -02  |                    | 22   |
| Methods of reporting to parents                | 26                                  | 12   | 23   | 06   | 15   | 08   | 10   | 20   |                    | 22   |
| Eigenvalues                                    | 12.62                               | 3.18 | 2.34 | 1.89 | 1.60 | 1.32 | 1.28 | 1.20 |                    |      |
| Percentage of total variance                   | 28.1                                | 7.1  | 5.2  | 4.2  | 3.5  | 2.9  | 2.8  | 2.7  |                    | 56.5 |
| Percentage of common variance                  | 56.8                                | 12.7 | 8.5  | 6.4  | 5.1  | 3.6  | 3.5  | 3.2  |                    | 100  |

*Age of the principals.* Four statistically significant correlation coefficients were obtained between the satisfaction scores and the age of the principals. The older the principals were, the more satisfied they tended to be with the factors of Resource Adequacy, Salary and Benefits, and Overall Satisfaction. On the other hand, the older principals tended to be less satisfied with Responsibility and Autonomy.

*Sex.* Only one statistically significant correlation was observed between sex and satisfaction scores. The male principals tended to be less satisfied with Salary and Benefits than were the female principals.

TABLE 3

STEPWISE MULTIPLE REGRESSION ANALYSIS WITH OVERALL SATISFACTION AS CRITERION AND THE FACTOR SCORES AS PREDICTORS WHEN CONTROLLING FOR SELECTED BACKGROUND VARIABLES OF PRINCIPALS (N=327)

| Variables Entered                  | R <sup>2</sup> | Increase in R <sup>2</sup> | r   |
|------------------------------------|----------------|----------------------------|-----|
| Background Variables               | .04            |                            |     |
| Responsibility and Autonomy        | .18*           | .14                        | .47 |
| Principal-Teacher Work Involvement | .26*           | .08                        | .42 |
| Liaison at District Level          | .30*           | .05                        | .37 |
| Status Recognition                 | .33*           | .03                        | .34 |
| Tasks Demands                      | .35*           | .02                        | .36 |
| Salary and Benefits                | .35            | .00                        | .22 |
| Resource Adequacy                  | .35            | .00                        | .24 |
| Rapport with Students              | .35            | .00                        | .19 |

\*Contribution significant at the .01 level

*Experience of the principals.* Three significant correlations were obtained between the satisfaction scores and the years of experience of the principals. The more experienced principals tended to be more satisfied with the factors Resource Adequacy and Task Demands, as well as with Overall Satisfaction.

*School size.* School size was significantly correlated with four of the factor scores. Principals in larger schools tended to be more satisfied with Resource Adequacy, Salary and Benefits, and Task Demands, and to be less satisfied with Rapport with Students than were principals of smaller schools.

*Location of schools (Urbanization).* Four statistically significant correlation coefficients were obtained between the satisfaction factor scores and the location of the schools, classified as rural, town or city. The more urbanized the school setting, the more satisfied principals were with Principal-Teacher Work Involvement, Resource Adequacy, and Salary and Benefits, and the less satisfied they were with Liaison at District Level.

*General Observations*

This study examined the job satisfaction of a sample of public school principals by means of a questionnaire containing 45 items obtained from the literature on job satisfaction. Three aspects of job satisfaction were examined: overall satisfaction, facet satisfaction, and satisfaction factors.

*Overall and Facet Satisfaction*

About 95 percent of the public school principals expressed overall satisfaction with their jobs. This satisfaction was also apparent in their responses to individual items, especially to items on relationships with teachers, students and the community. However, the principals' satisfaction was much lower on matters related to bargaining, counselling, use of time, and consultations with the board on working conditions.



TABLE 4

PEARSON CORRELATION COEFFICIENTS BETWEEN SATISFACTION FACTOR SCORES AND BACKGROUND VARIABLES

| Satisfaction Factors               | Background Variables |       |                           |            |        |   |
|------------------------------------|----------------------|-------|---------------------------|------------|--------|---|
|                                    | Educational Level    | Age   | Sex<br>1 Female<br>2 Male | Experience | Size   | Location<br>1 Rural<br>2 Town<br>3 City |
| Liaison at District Level          | -.04                 | .05   | .05                       | -.01       | .03    | -.21**                                  |
| Principal-Teacher Work Involvement | .03                  | .08   | -.04                      | .09        | -.07   | .12*                                    |
| Status Recognition                 | -.09                 | .06   | -.03                      | .03        | -.02   | .08                                     |
| Responsibility and Autonomy        | -.02                 | -.10* | -.04                      | -.08       | .02    | .01                                     |
| Resource Adequacy                  | .16**                | .23** | -.03                      | .21**      | .28**  | .34**                                   |
| Salaries and Benefits              | .02                  | .10*  | -.15**                    | .08        | .14**  | .20**                                   |
| Task Demands                       | .08                  | .00   | .10*                      | .16**      | .10*   | -.04                                    |
| Rapport with Students              | -.09                 | -.00  | .04                       | -.02       | -.16** | .06                                     |
| Overall Satisfaction               | -.03                 | .16** | -.01                      | .19**      | .08    | .03                                     |

\* p≤.05

\*\* p≤.01

The range of satisfaction with job facets was apparently not reflected in the level of Overall Satisfaction, indicating that overall satisfaction may be different from the sum of the satisfaction scores on distinct facets. The principals may be reasonably satisfied with their jobs, but at the same time dissatisfied with specific facets of their work.

Factor Scores

The varimax factor analysis yielded eight distinct factors of job satisfaction. Of these eight, five were significantly correlated with Overall Satisfaction, but not very substantially. In the regression analysis only 35 percent of the variance on Overall Satisfaction was accounted for by the factors. Responsibility and Autonomy was the best predictor of overall satisfaction, followed by Principal-Teacher Work Involvement, Liaison at District Level, Status Recognition, and Task Demands.

### *Background Variables and Job Satisfaction*

The age and experience of principals were the only background variables which were positively correlated with Overall Satisfaction. These two variables were also positively correlated with Resource Adequacy. In addition, age was positively correlated with Salary and Benefits, while experience was positively correlated with Task Demands.

The size of schools, though not significantly related to Overall Satisfaction, was correlated significantly with several satisfaction factors. Size was positively correlated with Resource Adequacy, Salary and Benefits, and Task Demands, and negatively correlated with Rapport with Students. Thus the larger the schools the more satisfied the principals tended to be with three of the satisfaction factors, and the less satisfied they tended to be with Rapport with Students.

The greater the urbanization of schools the more satisfied the principals tended to be with Resource Adequacy, and Salary and Benefits, and less satisfied with Liaison at District Level.

### *Discussion*

Several questions arise out of the findings on overall satisfaction of the school principals. For instance, how does the job satisfaction of principals compare with that of teachers? Are the principals perhaps too satisfied with their jobs? If so, why? Is the Overall Satisfaction measure an appropriate measure which can be used across different occupations? The low degree to which the factors predict Overall Satisfaction in this study suggest that overall satisfaction is achieved differently by different principals. For instance, one principal said, "My staff provides both great satisfaction and frustration." Another stated "Business management often consumes far more time than educational decisions on policy directions."

The identification of five significant predictors of overall satisfaction presents a further problem. Should other facets have been examined, or is overall satisfaction different from the sum of a series of satisfiers on a job? This study suggests that overall satisfaction is not necessarily the sum of the satisfiers on a number of items; in fact, several seemingly major factors in the job situation, such as Rapport with Students, appear not to be substantially correlated with overall satisfaction. Further study on the methodology of research on satisfaction is suggested by these findings.

The background variables in general were only minor contributors to the job satisfaction of the principals. Perhaps the two most significant findings in this area were that the principals in city schools generally had less satisfaction through Liaison at District Level, and that principals of larger schools had less satisfaction with Rapport with Students. These point to two areas where deliberate changes may be in order. In the first instance, the principals in urban schools sense a problem with their relationships with central office and/or with the superintendent. The communication between the two levels needs examination in order to reduce dissatisfaction related to this liaison.

Perhaps even more important is the finding that principals of larger schools tend to be more satisfied with Resource Adequacy and less with their Rapport with Students. Policies and practices need to be examined to see why the shift of these principals' satisfaction is away from students and toward resources.

A basic difference between principals and teachers is noted in regard to the sources of job satisfaction. According to Holdaway (1978), teachers' major source of

job satisfaction comes from their relationships with students. For the principals of this study this relationship was a minor though clearly identifiable factor in job satisfaction which, however, did not contribute to their overall satisfaction on the job. This falling away in importance of relationships with students as teachers become principals seems to increase as the schools become larger.

The identification of Resource Adequacy as a job satisfaction factor presents a different problem. It also is independent of the overall satisfaction of principals, yet increases in importance as a satisfier as schools become larger or more urbanized. Perhaps this change indicates a difference in the role of the principal between rural and urban schools. This finding supports the Herzberg (1959) claim that “the work itself” is a major factor in job satisfaction. Thus principals have less to do with students as schools increase in size and are more concerned with resource adequacy. The substantial correlation between Resource Adequacy and location of schools could be related to the current practices of decentralizing budget decisions to the school level.

Salary and Benefits, another job satisfaction factor identified, was positively related with the age of the principal, size of the school, and urbanization of the school; it was significantly related to the sex of the principals. Since it was not related significantly to Overall Satisfaction, these findings provide for interesting speculation. Older principals, principals in larger schools and city principals appear to be more satisfied with salary and benefits than the other principals; female principals tend to be more satisfied with salary and benefits than male principals.

The two major conclusions of this study were, first, that background variables did not contribute significantly to the Overall Satisfaction of school principals, and second, that the major contributors to Overall Satisfaction in order of importance were Responsibility and Autonomy, Principal-Teacher Work Involvement, Liaison at District Level, Status Recognition, and Task Demands. These factors illustrate the complexity of job satisfaction for administrators. Statements by the principals themselves illustrate the various job satisfaction factors in their work situation. Several examples are provided below.

*Responsibility and Autonomy.* “Having the authority and responsibility to effect changes in the educational process which are beneficial to both students and teachers.”

*Principal-Teacher Work Involvement.* “I enjoy the companionship and educational growth associating with staff personnel of the school.”

*Liaison at District Level.* “Lack of proper communication with Central Office—little or no consultation yet decisions made requiring a tremendous amount of administrative time.”

*Status Recognition.* “The relationship with all individuals, staff, parents, administration is usually positive and satisfying.”

*Task Demands.* “The additional little tasks that are expected by the Board, the teachers and students, and the community. There are not enough hours in a day.”

### *Conclusion*

The questions raised through this study emphasize the importance of research in the area of job satisfaction of administrators. The nature of satisfaction itself as it relates to the quality of work life needs further exploration. Whether the level of overall satisfaction expressed by principals or the range of levels of satisfaction on



individual facets is a better indicator of the quality of the working life of the principals is difficult to determine. Another aspect that should be examined is the relationship between the perception of levels of dissatisfaction and stress. Finally, further study is still needed to examine the importance of satisfaction for the effectiveness of the principal and more importantly for the effectiveness of the school, and to interpret the importance of all of the satisfaction variables for the quality of the working life of students and teachers, as well as principals.

### References

- Cochran, R. F. (1977). California school superintendents like their jobs. *Phi Delta Kappan*, 58, 777.
- Gruneberg, M. M. (1976). *Job satisfaction—A reader*. New York: Wiley.
- Herzberg, F., Mausner, B., & Snyderman, B. B. (1959). *The motivation to work*. New York: Wiley.
- Holdaway, E. A. (1978). Facet and overall satisfaction of teachers. *Educational Administration Quarterly*, 14, 30-47.
- Iannone, R. (1973). What motivates principals? *Journal of Educational Research*, 66, 206-262.
- Johnson, D. A., & Weiss, D. J. (1971). *Middle management, decision-making and job satisfaction*. Danville, IL: Interstate.
- Kerlinger, F. N. (1973). *Foundations of behavioral research*. New York: Holt, Rinehart & Winston.
- Lawler, E. E. (1973). *Motivation in work organizations*. Belmont, CA: Wadsworth.
- Locke, E. A. (1969). What is job satisfaction? *Organizational Behavior and Human Performance*, 4, 309-336.
- Locke, E. A. (1976). The nature and causes of job satisfaction. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 1296-1349). Chicago: Rand McNally.
- Maslow, A. B. (1970). *Motivation and personality*. New York: Harper & Row.
- Miner, J. B. (1980). *Theories of organizational behavior*. Hinsdale, IL: Dryden.
- Porter, L. W., & Lawler, E. E. (1968). *Managerial attitudes and performance*. Homewood, IL: Irwin.
- Porter, L. W., Lawler, E. E., & Hackman, J. R. (1975). *Behavior in organizations*. New York: McGraw-Hill.
- Rice, A. W. (1978). *Individual and work variables associated with principal job satisfaction*. Unpublished doctoral dissertation. Edmonton: The University of Alberta.
- Schmidt, G. L. (1976). Job satisfaction among secondary school administrators. *Educational Administration Quarterly*, 12, 68-86.
- Smith, P. C., Kendall, L. M., & Hulin, C. L. (1969). *The measurement of satisfaction in work and retirement*. Chicago: Rand McNally.
- Thronson, L. J. (1969). *The relationship of professional orientation of teachers to hierarchical authority, procedural specification and satisfaction*. Unpublished master's thesis. Edmonton: The University of Alberta.
- Van Maanen, J. P., & Katz, R. (1976). Individuals and their careers: Some temporal considerations for work satisfaction. *Personnel Psychology*, 29, 601-616.
- Vroom, V. (1964). *Work and motivation*. New York: Wiley.
- Wanous, J. P., & Lawler, E. E. (1972). Measurement and meaning of job satisfaction. *Journal of Applied Psychology*, 56, 95-105.

HERBERT J. WALBERG

*University of Illinois at Chicago*

RAM PAL SINGH

*Jialal Institute of Education, India*

and

SHIOW-LING TSAI

*University of Illinois at Chicago*

## The Achievement Productivity of Psychological Variables: A Replication and Extension in Rajasthan, India

*To determine the significance and relative influence of individual and environmental variables on learning, the achievement-test scores of 439 tenth-grade students in 39 schools in Rajasthan, India were regressed on IQ, socio-economic status (SES), self-rated motivation, peer-rated studiousness, peer-group status, and parent-rated home environment. The squared multiple-regression of .76 is highly significant; and both correlations and partial correlations conform closely in sign and magnitude to those that have been quantitatively synthesized from many studies. The hypothesized variables contribute uniquely to achievement variance; and three variables—SES, sex, and language medium of instruction—are superfluous in the context of the hypothesized variables.*

During the past two decades, several groups began research programs on students' perceptions of the psychological environments of their classroom social

---

Herbert J. Walberg holds a PhD in educational psychology and is Research Professor of Education at the University of Illinois at Chicago.

Ram Pal Singh holds a PhD in education and is Director of the Jialal Institute of Education, Rajasthan, India.

Shiow-Ling Tsai holds two master's degrees (education and economics) and is a doctoral student in public policy analysis at the University of Illinois at Chicago.

groups. Reviews (Randawa & Fu, 1973; Walberg & Moos, 1980) as well as a quantitative synthesis of twelve studies carried out on more than 17,000 students in four countries of English speakers (Haertel, Walbert, & Haertel, 1981) show that such perceptions obtained generally at mid-academic year account for substantial variance in cognitive, affective, and behavioral learning-outcome measures beyond that accounted for by corresponding beginning-of-course measures or mental abilities or both.

Thus, perceptions of the learning environment have incremental, predictive validity; and they are suggestive of the aspects of group climate or morale that promote various learning outcomes, some of which are rarely measured in conventional school testing programs and program evaluations. Such perceptions also have discriminant or criterion validity, that is, statistical sensitivity to educational and psychological conditions and treatments when used as criteria in experimental and correlational studies.

During the past decade, five investigations have extended or generalized the predictability of learning outcomes from classroom psychological environments to newly industrializing countries in the Caribbean, South America, and Central and South Asia (see Paige, 1979, for a detailed review). The sample characteristics and major findings of these studies are shown in Table 1.

TABLE 1  
CHARACTERISTICS OF FIVE STUDIES OF PSYCHOLOGICAL  
LEARNING ENVIRONMENTS IN DEVELOPING COUNTRIES

| Author                           | Sample   | <u>R<sup>2</sup> and Dependent Variables</u>              | <u>Generally Significant Independent Variables</u>   |
|----------------------------------|--|---|--|
| Holsinger (1973)                 | 2,533 students in 90 third, fourth & fifth grade classes in the Federal District of Brazil                             | .19, .64 modernity<br>.09, .52 information                | Amount of schooling, age, socioeconomic status, classroom environment, television exposure, school physical environment, peer group, and teacher attitudes; individual and class level R <sup>2</sup> 's are reported first and second in each row |
| Persand (1976)                   | 1,277 third and sixth grade students in 18 schools in Jamaica  | -social development<br>-aspirations                       | Student perceptions of school task orientation, work organization, psychological support, cooperation and innovation; school staff characteristics, parent and peer socioeconomic status; television exposure                                      |
| Walberg, Rasher and Singh (1977) | 166 groups of tenth grade students in 150 classes in 26 schools in Rajasthan, India                                    | .67 science achievement<br>.81 school studies achievement | IQ, amount of Teacher training, and classroom psychological environment were significant in regressions when controlled for one another and other variables  |
| Paige (1977)                     | 1,621 sixth grade students in 30 urban and 30 rural schools in East Java, Indonesia                                    | .12 modernity<br>.27 achievement                          | Pupil background characteristics, home learning environment, school and class physical environment and organizational characteristics, and school location   |
| Present Study (1982)             | 439 high and low achieving tenth grade students in 39 predominantly high and low achieving schools in Rajasthan, India | .76 achievement   | IQ, classroom environment, home environment, peer group, socioeconomic status, and pupil studiousness  |

Several studies in Table 1 complement those from more industrialized countries because they extend the learning criteria beyond the usual outcomes and demon-



strate the cross-cultural validity of environment measures. Holsinger (1973), for example, examined a variety of environmental influences in Brazil on “modernity,” the tendency for individuals to defer immediate gratification for longterm benefits, to plan small families, and to engage in similar behavior conducive to individual and national success in developing countries. Student scores on modernity scales and information tests were found to be significantly related individually and collectively to several sets of student and environmental variables, as indicated in Table 1. The studies in Table 1 statistically control the environment-learning correlation for plausible determinants of learning other than IQ and prior achievement. Several of the studies, moreover, have an advantage over typical ones conducted in the United States in that achievement tests geared specifically to required national or provincial curricula are employed, which removes a large source of error variation attributable to differences in exposure to subject matter and the mismatch of measured outcomes and educational intents.

### *General Influences on Learning*

Such studies avoid the patently unsustainable assumption that the classroom is a closed system. They instead assume that factors outside the school such as the home and peer-group environments influence student ability and motivation as well as classroom climate, responsiveness to instruction, and achievement performance, and further that learning itself takes place to a large extent outside the class. The possible importance of outside influences can be seen in the relatively small percentage of waking potentially-educative hours from birth through age 18 that the student spends in school—about 14.1—assuming 10 hours of sleep, 180 six-hour school days per year, and perfect attendance. More comprehensive studies that include both fixed variables such as IQ and socio-economic status as well as potentially alterable variables in the class, home, and peer group afford estimates of the relative importance of each variable as a part of a larger set.

More can be learned from such comprehensive studies that include larger sets of such variables than from the usual bivariate or single covariate study. Multiple regressions, moreover, that simultaneously control the variables for one another provide more stringent tests of causality and practical suggestions for improvements in educational productivity within and outside the schools.

From quantitative syntheses of hundreds of studies, it appears that school learning depends on nine proximal psychological constructs: student age or developmental level, motivation, and ability or prior achievement; quantity and quality of instruction (including self-instruction); the social-psychological environments of the home, classroom group, and peer group; and exposure to mass media, particularly television (*Research Synthesis: The State of the Art*, 1980; Walberg, 1981). Although the studies of Table 1 neglect quantity and quality of instruction and mass media exposure on the whole, they include from three to five of the constructs; and, in some cases, they account for a considerable amount of variance in learning outcomes. The purpose of the present study is to replicate the earlier study in Rajasthan and to probe simultaneously the validity of five of the constructs—student ability and motivation, and classroom, home, and peer-group environments.

### *Method*

#### *Sample*

Both purposiveness and economy determined the sampling plan. It was hoped that a relatively small set of schools and students could be sampled that would vary

greatly in achievement outcomes. Larger variation in outcomes, of course, raises the possibility of detecting significant covariation with alterable independent variables and allows interpolation rather than extrapolation of conclusions over a wider range.

From 239 schools in the State of Rajasthan, India that participated in the tenth-grade Board of Secondary Examinations program conducted by government, the two sets of fifteen schools with the best and worst average scores on objective tests geared to the standard curriculum were selected. In addition, to obtain a somewhat larger sample than 30 schools and to include some middle-scoring schools conveniently, all nine schools within the Central Political Territory were included, which resulted in a sample of 39 schools.

From each school, the students with the top five and bottom five scores on the examinations were selected for participation in the study. In several schools where absences were anticipated because of holidays, a few more than the ten extreme-scoring students were included. For this reason, the total sample included 439 instead of 390 students.

### *Variables and Testing Procedure*

Tables 2 through 4 describe the psychometric and other variables used in the study. Soon after the achievement tests were scored and analyzed, the students completed the IQ, classroom environment, and motivation scales in school; and the parents filled out the home-environment questionnaire in their homes.

TABLE 2

#### VARIABLE DESCRIPTION, SAMPLE ITEMS AND UNIVARIATE STATISTICS

| <u>Variables</u>                  | <u>Description or Sample Items</u>  | <u>Mean</u> | <u>Standard<br/>Deviation</u> |
|-----------------------------------|---|-------------|-------------------------------|
| Students' achievement             | The percentage of correct answers to extensive multiple-choice achievement items secured on the Secondary School Examination in science, social studies, language, and mathematics conducted by the Board of Secondary Education, Rajasthan, India. Internal consistency reliabilities of components in previous analyzed average in the .70s. "From the point of view of raw materials, which is the best place for establishing a sugar factory--Ajmer, Bombay, Lucknow, or Kampur?" "Which is the best method for prevention of cholera--Vaccination, Innoculations, or Saline Injection?" | 54.23       | 12.54                         |
| Pupil's IQ                        | The IQ test (in Hindi) with 100 short items on number series, deduction, vocabulary, synonyms and antonyms, analogies, and general knowledge; an internal consistency reliability was .94 in a large, heterogenous sample of Rajasthan.   | 96.40       | 12.18                         |
| Motivation rated by students      | 23 items on the questionnaire:<br>"At home, when I start a new project, I usually Finish it...5...4...3...2...1...Don't finish it."<br>"When I do something well, it is because I Work hard...5...4...3...2...1...Was lucky."<br>"I am proud of my school work<br>All the time...5...4...3...2...1...Never."  | 72.44       | 15.90                         |
| Home environment rated by parents | 45 items on the questionnaire:<br>"Does your child have a library card? Yes <input type="checkbox"/> No <input type="checkbox"/> "<br>"Does the child usually watch television without parent there? Yes <input type="checkbox"/> No <input type="checkbox"/> "<br>"Do you know your child's best subject in school? Yes <input type="checkbox"/> No <input type="checkbox"/> "   | 83.89       | 23.38                         |

TABLE 3

SCALES, SAMPLE ITEMS, RELIABILITIES, AND STATISTICS  
OF FIFTEEN SUBSCALES FROM THE LEARNING ENVIRONMENT INVENTORY

| <u>Scales</u>        | <u>Sample Item</u>                                      | <u>Alpha<br/>Reliability</u> | <u>Mean</u> | <u>Standard<br/>Deviation</u> |
|----------------------|---|------------------------------|-------------|-------------------------------|
| Cohesiveness         | All students know each other well.                      | .69                          | 11.68       | 3.19                          |
| Satisfaction         | The students enjoy their classwork.                     | .79                          | 10.82       | 3.29                          |
| -Friction            | Certain students are responsible for petty quarrels.    | .72                          | 12.89       | 5.03                          |
| -Apathy              | Students do not care what the class does.               | .82                          | 4.27        | 4.93                          |
| -Cliqueness          | Some students refuse to mix with the rest of the class. | .65                          | 8.48        | 3.26                          |
| Democracy            | Students have about equal influence on the class.       | .67                          | 6.43        | 4.50                          |
| -Diversity           | Interests vary greatly within the group.                | .53                          | 14.31       | 5.92                          |
| -Favoritism          | Certain students are favored more than others.          | .78                          | 10.23       | 4.01                          |
| Goal Direction       | Each student knows the goals of the course.             | .85                          | 10.05       | 3.44                          |
| Material environment | The room is bright and comfortable.                     | .56                          | 13.57       | 4.13                          |
| Formality            | The class has rules to guide its activities.            | .76                          | 13.61       | 4.32                          |
| Difficulty           | Students find the work hard to do.                      | .64                          | 6.11        | 4.28                          |
| Speed                | The course material is covered quickly.                 | .70                          | 9.30        | 3.07                          |
| -Disorganization     | The class is inefficient and not organized.             | .82                          | 8.16        | 4.19                          |
| Competitiveness      | There is much competition in the class.                 | .78                          | 9.35        | 3.43                          |

NOTE: Subscales preceded by a minus sign measure aspects of the sociopsychological environment that are usually negatively related to learning outcomes.

TABLE 4

VARIABLES, CODING AND PERCENTAGES

| <u>Variables</u>     | <u>Coding and Percentages</u>   |
|----------------------|---|
| Peer-group           | Government school = 1 (54.5%)<br>Private, tuition school = 2 (45.5%)  |
| Female students      | Boys = 1 (69.1%)<br>Girls = 2 (30.9%)   |
| Socioeconomic status | Graded by three dimensions--<br>education, occupation and income of parents:<br>Average = 1 (39.5%)<br>High = 2 (60.5%) |
| English instruction  | Hindi = 1 (63.6%)<br>English = 2 (36.4%)  |
| Studiosness          | Classroom sociometric rating of pupils'<br>classroom behavior:<br>Mischievous = 1 (16.4%)<br>Studiosness = 2 (83.6%)    |

Results and Discussion

As shown in Table 5, the achievement scores of the students were regressed on seven combinations of three sets of the independent variables—IQ, classroom environment, and other variables. The variances accounted for by all combinations of sets are significant beyond the .001 level. The coefficient of determination of the regression equation containing all the variables, .76, corresponds to a multiple correlation of .87 which exceeds the approximate reliability of the achievement measure.



TABLE 5  
REGRESSIONS OF ACHIEVEMENT ON SETS OF VARIABLES

| <u>Sets of Independent Variables</u> | <u>Number of Variables</u> | <u>R<sup>2</sup></u> |
|--------------------------------------|----------------------------|----------------------|
| Student IQ                           | 1                          | .620                 |
| Learning environment                 | 15                         | .498                 |
| Other variables                      | 9                          | .550                 |
| IQ and environment                   | 16                         | .702                 |
| IQ and other variables               | 10                         | .707                 |
| Environment and other variables      | 24                         | .679                 |
| IQ, environment and other variables  | 25                         | .758                 |

NOTE: All R<sup>2</sup> are significant at the .001 level; and the incremental variances attributable to adding environment, the other variables, and both to the regression containing IQ are also significant at the .001 level.

IQ, the sets of classroom environment variables, and the set of other variables each account for about half or slightly more of the variance in achievement. Moreover, the addition of the learning environment variables, the other variables, or both sets produces significant increments in the variance beyond that accounted for by IQ; the F-ratios are respectively 7.76, 14.15, and 9.81, all of which are significant beyond the .001 level. By subtraction of the variances in Table 5, IQ uniquely adds 7.9 percent to the explained variance when entered last, class environment similarly adds 5.1 percent, and the other variables add 5.6 percent.

#### *Specific Variables*

The first column in Table 6 shows the 95 percent confidence intervals of previous estimates of achievement correlations with the independent variables for those that are available (*Research Synthesis: The State of the Art*, 1980). These intervals are the mean correlations plus and minus two standard deviations or the fitted values from analyses of variance or regressions plus and minus two standard errors of estimate.

The correlations in the present study correspond reasonably well in sign and magnitude to those revealed by past research. As shown in Table 6, achievement correlated positively with Classroom Formality, the Material Environment, Cohesiveness, Satisfaction, Goal Direction, Speed, Competitiveness, and Diversity, and negatively with Apathy, Favoritism, Cliqueness, and Disorganization. Achievement also correlated positively with Peer Group, Socio-Economic Status, Studiousness, Motivation, IQ, and Home Environment. The correlations are generally higher than those in past research probably because of the wide achievement variation in the sample.

Sex, Motivation, Socio-Economic Status, and English as the medium of instruction are insignificant when the variables are controlled for one another by partial correlation, because they are weakly correlated with achievement, because they are correlated with the other independent variables, or both. The variables that remain

significant in partial correlation analysis are the classroom environment set, Peer Group, Studiousness, IQ, and Home Environment.

TABLE 6  
CORRELATIONS AND PARTIAL CORRELATIONS WITH STUDENTS' ACHIEVEMENT

|                      | Prior Estimate | Present Sample |         |
|----------------------|----------------|----------------|---------|
|                      |                | Estimates      |         |
|                      | Pearson        | Pearson        | Partial |
| Learning Environment |                |                |         |
| Formality            | (-.17)- .61    | .62**          | .19**   |
| Material Environment | .02 - .80      | .49**          | .25**   |
| Cohesiveness         | .07 - .85      | .39**          | .16**   |
| Satisfaction         | .09 - .87      | .32**          | .16**   |
| Goal Direction       | .05 - .83      | .30**          | .18**   |
| Speed                | (-.29)- .49    | .22**          | .14**   |
| Competitiveness      | (-.29)- .49    | .17**          | .11     |
| Diversity            | (-.29)- .49    | .15            | -.01*   |
| Friction             | (-.86)-(-.08)  | -.02           | -.09    |
| Democracy            | .03 - .81      | -.05           | .01     |
| Difficulty           | .00 - .78      | -.09**         | -.03**  |
| Apathy               | .18 -(-.60)    | -.30**         | -.13**  |
| Favoritism           | .08 -(-.70)    | -.31**         | -.25**  |
| Cliqueness           | .12 -(-.66)    | -.38**         | -.20**  |
| Disorganization      | .13 -(-.65)    | -.50**         | -.25**  |
| Peer Group           | .20 - .36      | .20**          | .15**   |
| Female Student       | ---            | .25**          | .07     |
| Socioeconomic Status | (-.14)- .78    | .40**          | .02     |
| English Instruction  | ---            | .43**          | .003    |
| Pupils' Studiousness | ---            | .51**          | .15**   |
| IQ                   | .57 - .87      | .79**          | .50**   |
| Motivation           | .25 - .41      | .50**          | .08     |
| Home Environment     | .36 - .76      | .69**          | .30**   |

NOTE: The partial correlations are controlled for all independent variables except in the case of the learning environment variables which are not controlled for one another.

\* = .05  
\*\* = .01

Conclusion

Contrary to some recent writings (see, for example, McDermott's [1976] edited collection, *Indeterminacy in Education*), quantitative syntheses of prior psychological research on classroom learning show fairly consistent, robust, and strong findings. It seems apparent, however, that learning is determined or influenced by multiple factors, perhaps nine, of which five proved significant when controlled for one another in the present study. Moreover, other factors, included for comparative purposes, including Socio-Economic Status and sex appear to be mediated by the more proximate factors. Quantity and quality of instruction, age or developmental level, and exposure to mass media were not included among the factors in this study; and it would be valuable to incorporate them in future investigations.

The importance of the present study is that it shows the generality of the significance of the factors in a culture far different from that of Australia, Canada and the United States. Along with related research, it suggests a valuable generality or universality of results concerning the consistent influential factors in learning.

#### References

- Haertel, G. D., Walberg, H. J., Haertel, E. H. (1981). Social-psychological environments and learning: A quantitative synthesis. *British Educational Research Journal*, 7, 27-36.
- Holsinger, D. B. (1973). The elementary school as modernizer. *International Journal of Comparative Sociology*, 14, 180-202.
- McDermott, J. C. (1976). *Indeterminacy in education*. Berkeley, CA: McCutchan.
- Paige, R. M. (1979). Learning environments in developing countries. In R. H. Moos (Ed.), *Evaluating educational environments*. San Francisco: Jossey-Bass Publishers.
- Persuad, G. (1976). School authority pattern and students' social development in selected primary schools in Jamaica. Unpublished PhD dissertation, Stanford University.
- Randawa, B. S., & Fu, L. L. W. (1973). Assessment and effect of some classroom variables. *Review of Educational Research*, 43, 303-321.
- Research Synthesis: State of the Art, Evaluation in Education*. (1980). Special issue on research synthesis. 4, 1-144.
- Walberg, H. J. (1981). A psychological theory of educational productivity. In F. H. Farley, & N. Gordon (Eds.), *Psychology and education*. Berkeley, CA: McCutchan.
- Walberg, H. J., & Moos, R. M. (1980). Assessing educational environments. *New Directions in Testing and Measurement*, 7, 63-76.
- Walberg, H. J., Singh, R. P., & Rasher, S. P. (1977). Predictive validity of student perception: A cross-cultural replication. *American Educational Research Journal*, 14, 45-49.



FLORENCE J. MAYNES

and

JOHN A. ROSS

OISE Trent Valley Centre

## Integrating School Curricula with Social Concerns

*The attempt to integrate school curricula with social concerns has been a recurrent motif in North American education: the social data presented to students, or which is collected by them, is intended to contribute to a greater understanding of major social themes and the resolution of socially relevant problems apposite to each theme. In this article impediments to the implementation of the progressive ideal are outlined, and it is proposed that these obstacles can be reduced by incorporating a cognitive approach to social issues analysis. Evidence on the effectiveness of the approach is presented from a delayed treatment instructional experiment involving 320 students in grades 7 to 10 in one Ontario school system, with additional data provided by a partial replication involving 159 grade 7 students in a second Ontario school system.*

The attempt to integrate school curricula with social concerns is a recurrent motif in North American education. Its essence is the belief that broad social themes should comprise a major portion of the school agenda and should be explicitly addressed in problem or issue contexts. In this approach the social data presented to students, or which is collected by them, is intended to contribute to a greater understanding of major social themes and the resolution of socially relevant problems apposite to each theme. It is anticipated that the information assimilated by students will be actively integrated into their real life concerns and lead ultimately to an improvement of society as these enlightened students graduate as

---

Florence Maynes (BA, University of Alberta; MA, University of Toronto) is a Research Associate with the OISE Trent Valley Centre, with research interests in instructional improvements and intrinsic motivation.

John Ross (PhD, McMaster University) is Associate Professor of Curriculum and Head of the OISE Trent Valley Centre. His major research interests are instructional design and program evaluation.

responsible citizens. The leading contemporary exponents of this approach are Bourne and Eisenberg (1978), Manson and Vuicich (1977), Newmann (1975), Oliver and Newmann (1967-1973), and Oliver and Shaver (1974).

### *Problem Statement*

Implementation of this ideal has been episodic, with the familiar pattern shared with other innovations of initial enthusiasm followed by a gradual return to mainstream approaches: the cross-disciplinary curriculum fights a losing battle against the core subjects orientation, and discussion classes aimed at the resolution of social issues struggle vainly against the twin threats of student apathy and disorder. Enthusiasts continue to exhort teachers to take up the movement. Why are these intuitively sensible approaches not adopted in more classrooms?

One set of explanations has to do with the relationship between schools and their social milieu, a relationship which is best understood in historical perspective. Historically the attempt to integrate curricula and social concerns has waxed and waned with the fate of the much more inclusive progressive education movement. Progressivism in education reached its philosophical heights in the first two decades of this century, became conventional wisdom in the depression and war years, experienced a severe decline under ferocious attack in the 1950s, revived unexpectedly to flourish in the turbulent 1960s and continued as a vibrant minority position in the back-to-basics era of the 1970s and 1980s.

Progressive education embraced an enormous diversity of ideals, practices and beliefs about education: identifying the core of the movement is fraught with peril. The most authoritative account (Cremin, 1962) categorically asserts that a capsule definition of the movement can never be devised because of its pluralistic and frequently contradictory character. Graham (1967), in reviewing the organizational development of progressivism, suggests that the phrase progressive education has been a normative rather than descriptive term, representing everything that was good about education in the 1920s and everything that was bad in the 1950s. Ravitch (1983) finds there has never been a clear cut definition of progressive education and that there has been more agreement about what it is not than what it is. The ideals of the movement were often honored in the breach, idiosyncratically applied and caricatured beyond recognition. These caveats notwithstanding, a catalogue of progressive ideals can be compiled:

- The curriculum should be designed in terms of the perceived needs of students; curriculum topics should be prioritized in terms of expected utility to the student and expanded to include health, vocational, family and community concerns.
- Instructional practices should be based on empirical research, which implies among other things that traditional subjects should be discarded in favor of a core curriculum focused on themes derived from the student's immediate experience, and that students should learn through direct experience rather than through passive encounters with texts.
- Schools should provide education to all types of children; instruction should be tailored to individual differences in their interests and abilities, with the emphasis on individual development rather than on the acquisition of knowledge.
- Instructional activities should be planned cooperatively by students and teachers; they should promote cooperation among students rather than competition and should provide ample opportunity for students' self-expression.



- Finally, the school should be closely linked to the community; for some this meant that the school should play a pivotal role in the reform of the community—for others it meant that students should learn to adjust to the existing community.

Progressivism in education emerged as part of the political protest movement of the same name in the first quarter of the century: it was the educational arm of the political movement. But in the subsequent decade it lost touch with its social and political foundations and became isolated as an organization limited to professional educators. Progressive education arose in response to conditions in the larger society and it fell when these conditions changed and it did not. The movement failed to keep pace with the transformation of society (Cremin, 1962); by not responding to new social issues its slogans became ossified and devoid of relevance.

Historical review of the progressive movement suggests that to be successful, attempts to integrate school curricula with social concerns must accurately appraise the society in which schools function and be responsive to it. But despite the importance of relating appropriately to the milieu, it is not sufficient. For example, even at the height of progressive influence there were constant reports of resistance on the part of teachers to the social concerns thrust, and evidence (recounted by Cremin, 1962, and Ravitch, 1983) that the task of teaching in this approach exceeded the competencies of a significant proportion of the teacher population. This suggests that a second set of explanations, focused on classroom variables and complementary to milieu factors, might assist in understanding why attempts to integrate school curricula with social concerns are less frequently adopted than one would expect:

1. There are distinctly different orientations to curriculum consisting of integrated belief systems about the ultimate ends of education and the means to accomplish these ends (Eisner & Vallance, 1974; Roberts, 1982; Miller, 1983). The social concerns thrust is simply one of many competing conceptions that may be less congruent with teacher preferences than alternative positions. This argument would help explain the reluctance of many; it would not account for those who express initial interest but who are unable or unwilling to follow through on their intentions.

2. Available classroom resources are typically subject oriented, aimed at the transmission of the content of a discipline. Efforts to break out of these boundaries require resourceful teachers able and willing to track down articles, pamphlets and fugitive documents. Although a few specific curriculum products have been produced for teachers at various times (e.g., Rugg's textbook series in the 1920s and '30s, the Harvard Public Issues Series of the late 1960s and the OISE Canadian Critical Issues Series of the 1970s), these must be constantly revised. So too, teachers may assess their own expertise in social issues pessimistically and be uncomfortable when the need to broach these limits is perceived.

3. Teachers may be uncertain as to the specifics of what they are to accomplish. The themes explored appear to have limitless scope. What are the limits/goals of in-class investigation? What are the anticipated knowledge objectives of social problem-solving?

4. Equally important may be uncertainties about the means to be employed by students in making sense of social questions, the intellectual skills they should and must acquire to deal rationally with social problems and issues. The appealing image of a classroom of eager young people sharing and discussing the results of their investigations and arriving at a resolution of the problem dissolves in the reality of twenty or thirty individuals each clamoring to have his or her opinion heard. Stu-



dents do not necessarily become mature problem-solvers without expert guidance and assistance; e.g., a recent study (Ross, 1981) shows little or no improvement from grade 7 to grade 10 in the strategies developed by students to solve personal decision problems.

5. Uncertainties about means and ends are likely to create apparently insoluble assessment problems. Teachers may be reluctant to attempt social issues analysis if the standards for judging the product of the analysis are unknown or ambiguous. In the absence of clear, specific and measurable standards, teachers have a soft and limited evidentiary base for assessing student performance and for judging the effectiveness of instructional efforts.

### *A Cognitive Approach*

The preceding list of problems constitutes a formidable set of obstacles to widespread adoption of the social concerns thrust and it would be unrealistic to seek panaceas. Potential solutions to some of these problems can be derived from recent developments in cognitive approaches to curriculum design. By cognitive approaches we mean orientations to instruction that focus on the overt teaching of the covert mental operations that underlie intellectual performance. One can of course view cognitive approaches as a competing platform of curriculum beliefs which is antithetical to the social concerns thrust, but it is helpful to consider the possibilities of integrating a cognitive orientation into a curriculum concerned with social issues. It should be noted that cognitive psychologists (e.g., Brown, 1978; Broudy, 1977; Spiro, 1977) recognize the importance of integrating school and out of school learning, even though some tend to treat the integrating process as more of a problem impeding school learning than as a goal of instruction (e.g., Champagne & Klopfer, 1981; Clement, 1981).

This article outlines the development and pilot testing of an instructional program that attempts to reduce some of the classroom obstacles previously listed which impede the implementation of the social concerns thrust. The study was undertaken as one component of a many-faceted set of programs to improve the intellectual performance of students in elementary and secondary schools. (An account of the scope of this set of programs is provided in Ross and Maynes, 1983b.)

Students who attempt to analyze social issues are required to perform a number of interrelated tasks of varying complexity. The core task involves decision-making: the synthesis of social information to select the best course of action from an array of competing alternatives. This core task can be decomposed into a set of component skills. First and foremost among these is the ability to focus the decision problem and bring it within manageable scope, and to develop a plan, or framework, within which the solution may be sought. Other skills follow from these, for example locating appropriate data, assessing its adequacy and dealing with conflicts and variability among data sources, selecting, recording and interpreting relevant data, drawing a valid conclusion and finally reporting the findings in a coherent way. Performance of the latter set of skills, however, is directed and shaped by the quality of the overall framework that the decision-maker has developed. Consequently, in this project emphasis was placed on designing instruction that would help students formulate a focus and framework for decision problems.

The second major emphasis was placed on the development of growth schemes (or learning hierarchies) for these skills. Mature, literate adults focus and frame their approach to social decisions in a sophisticated way. Their behavior can be

analyzed and described. Very young children, on the other hand, approach their decisions in a very unsophisticated manner, which can also be analyzed and described. Between the two exists a gap which is variously bridged by experience and practice and which leads to skill maturation and the accumulation of appropriate content knowledge. It follows then that those who would improve the quality of children's experience through the medium of instruction should be aware that the mature performance may be initially beyond the grasp of most elementary or even early secondary school children. However, if incremental levels between the upper and lower extremes are described and instruction matched to these, students can be helped towards the upper goal by focusing successively on the lower and intermediate levels. In this way each identified skill increment can be achieved without overloading the capacity of the students, and teachers may temporarily discontinue input at the point at which feedback alerts them to the need for extensive practice and consolidation before proceeding further.

Growth schemes for the two core skills of the core task are given in Tables 1 and 2. In each case the main dimensions of growth involve the explicitness of the decision problem, the number of alternative courses of action considered and the explicitness of the criteria used to assess alternatives.

TABLE 1  
GROWTH SCHEME FOR THE SKILL OF  
FOCUSING A DECISION-MAKING PROBLEM

| Level | Level Characteristics  | Example<br>(Issue: Canada's energy resources)  |
|-------|--|--|
| 1     | factual question not calling for decision  | How much oil does Canada have?   |
| 2     | decision question/alternative courses of action and criteria not identified                  | What should Canada do about the energy situation?  |
| 3     | one course of action identified/criteria for judging it not specified                        | Should Canada get its future energy supplies from coal?                                      |
| 4     | two or more alternative courses of action identified/criteria for judging them not specified | Is coal a better energy source than oil?   |
| 5     | two or more alternative courses of action identified/one or more explicit criteria specified | Would it be safer and cheaper for Canada to depend on solar power rather than nuclear power? |

These growth schemes became the blueprints for the instructional sequences that were piloted in the study. Since it was anticipated that students would have had little or no previous experience with formal decision-making procedures, the lessons began with the lowest level of the growth scheme. A student-teacher discussion format was used. As each successive level was introduced, its advantages relative to the previous level were pointed out through specific examples. Further examples were then provided and the students began to take the initiative for devel-

oping the focus and framework. Finally, additional examples for individual student practice were supplied. Evaluation of student performance with these exercise examples allowed teachers to diagnose and remediate problems of individual students. Additional detail on the instructional design process is provided in Ross and Maynes (1983b).

TABLE 2  
GROWTH SCHEME FOR THE SKILL OF ESTABLISHING  
A FRAMEWORK FOR A DECISION-MAKING PROBLEM

| Level  | Level Characteristics  | Example (Where should the shopping mall be built?)  |                                |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
|--|--|---|--------------------------------|--|-------------------------|---|--|--|---------------------------------|---|---|---|---------------|---|--|---|-----------------------------|---|---|---|-------------------------|---|---|---|
| 1  | one course of action/<br>feeble justification<br>given for it  | <table><tr><td>Build Mall Downtown</td></tr><tr><td>near our school so I<br/>can go to MacDonald's</td></tr></table>  |                                |  | Build Mall Downtown     | near our school so I<br>can go to MacDonald's |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| Build Mall Downtown  |  |   |                                |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| near our school so I<br>can go to MacDonald's  |  |   |                                |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| 2  | several alternative<br>courses of action/<br>different advantages<br>identified for each<br>alternative                  | <table><tr><td>Build the Mall Downtown</td><td>Build the Mall East of City</td></tr><tr><td>-lots of people already<br/>go downtown<br/>-good roads<br/>-downtown merchants want<br/>it here</td><td>-good flat land<br/>-no buildings would have to be<br/>torn down<br/>-people in east end would have<br/>have a place to shop</td></tr></table>   |                                |  | Build the Mall Downtown | Build the Mall East of City                   | -lots of people already<br>go downtown<br>-good roads<br>-downtown merchants want<br>it here | -good flat land<br>-no buildings would have to be<br>torn down<br>-people in east end would have<br>have a place to shop |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| Build the Mall Downtown  | Build the Mall East of City  |   |                                |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| -lots of people already<br>go downtown<br>-good roads<br>-downtown merchants want<br>it here | -good flat land<br>-no buildings would have to be<br>torn down<br>-people in east end would have<br>have a place to shop |   |                                |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| 3  | several alternative courses<br>of action/different<br>advantages and<br>disadvantages identified<br>for each alternative | <table><tr><td></td><td>Build the Mall Downtown</td><td colspan="2">Build the Mall East of City</td></tr><tr><td>Advantages</td><td>-lots of people available<br/>-good roads<br/>-downtown merchants want<br/>it here</td><td colspan="2">-good flat land<br/>-no buildings would have to<br/>be torn down<br/>-would give people in east<br/>end a place to shop</td></tr><tr><td>Disadvantages</td><td>-would have to move some<br/>stores<br/>-land is very expensive<br/>-requires bylaw change</td><td colspan="2">-would have to build new<br/>roads<br/>-farmland would be lost</td></tr></table> |                                |  |                         | Build the Mall Downtown                       | Build the Mall East of City  |  | Advantages                      | -lots of people available<br>-good roads<br>-downtown merchants want<br>it here | -good flat land<br>-no buildings would have to<br>be torn down<br>-would give people in east<br>end a place to shop |   | Disadvantages | -would have to move some<br>stores<br>-land is very expensive<br>-requires bylaw change | -would have to build new<br>roads<br>-farmland would be lost |   |                             |   |   |   |                         |   |   |   |
|  | Build the Mall Downtown  | Build the Mall East of City   |                                |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| Advantages   | -lots of people available<br>-good roads<br>-downtown merchants want<br>it here  | -good flat land<br>-no buildings would have to<br>be torn down<br>-would give people in east<br>end a place to shop   |                                |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| Disadvantages  | -would have to move some<br>stores<br>-land is very expensive<br>-requires bylaw change                                  | -would have to build new<br>roads<br>-farmland would be lost  |                                |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| 4  | several alternative courses<br>of action/explicit criteria<br>systematically applied to<br>each                          | <table><tr><td></td><td>Build the Mall<br/>Downtown</td><td>Build the Mall<br/>East of City</td><td>Build the Mall<br/>West of City</td></tr><tr><td>What downtown<br/>merchants want</td><td>+</td><td>-</td><td>-</td></tr><tr><td>Cost of land</td><td>-</td><td>+</td><td>+</td></tr><tr><td>Availability of<br/>shoppers</td><td>+</td><td>+</td><td>-</td></tr><tr><td>Suitability of<br/>roads</td><td>+</td><td>-</td><td>-</td></tr></table>   |                                |  |                         | Build the Mall<br>Downtown                    | Build the Mall<br>East of City   | Build the Mall<br>West of City   | What downtown<br>merchants want | +   | -   | - | Cost of land  | -   | +  | + | Availability of<br>shoppers | + | + | - | Suitability of<br>roads | + | - | - |
|  | Build the Mall<br>Downtown   | Build the Mall<br>East of City  | Build the Mall<br>West of City |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| What downtown<br>merchants want  | +  | -   | -                              |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| Cost of land   | -  | +   | +                              |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| Availability of<br>shoppers  | +  | +   | -                              |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |
| Suitability of<br>roads  | +  | -   | -                              |  |                         |   |  |  |                                 |   |   |   |               |   |  |   |                             |   |   |   |                         |   |   |   |

The context for the pilot testing of the instructional intervention was Intermediate Geography. Two versions of the program were prepared based on the content requirements of the Ontario Ministry of Education Guidelines for Geography. One unit was based on the Population Patterns section of the Guideline (taught in grade 7 in the pilot school system) and the other was based on the Canadian Resources section (taught in grades 9 or 10 in the pilot school system). A large excerpt of the grades 9 to 10 version is provided in Ross and Maynes (1982, chapter 4).

Evaluation Materials

The growth schemes also provided blueprints for the creation of multiple-choice test items to assess student achievement of the skills. Each of the response alternatives on a single item presented a framework or a focusing question, as the case



might be, that illustrated a level of the growth scheme. The task of the respondent then was to select the best response option; the assumption was that a student would recognize and select the response that matched his or her understanding of the way the problem should be approached.

Each focusing item contained five alternatives representing levels 1 to 5 of Table 1. Each frameworking item contained four alternatives, one corresponding to each of levels 1 to 4 of Table 2. A pool of such test items was available (Ross & Maynes, 1981). Six for each skill were selected and two of each were randomly assigned to one of three test forms, A, B and C.

An additional ten-point item for assessing the students' ability to select the appropriate data from a chart, graph or article and transfer it to the appropriate cell of an alternatives-by-criteria framework representation (i.e., level 4 of Table 2) was included in each booklet.

Each booklet contained two additional open-ended test items. In each a situation, similar to the situations in the multiple-choice items, was presented and students were required to produce rather than select a response, either the focus or the framework they would apply to the problem. It was anticipated that student performance on the open-ended items would be lower (because the task is more demanding) but that the pattern of results for the constructed and multiple-choice items would be similar (Ausubel & Robinson, 1969; Ebel, 1979; Ross & Maynes, 1983a).

### *Evaluation Design*

A delayed treatment experiment was designed. Fourteen volunteer classrooms in one school system were identified. Classrooms were randomly assigned to the early treatment condition (i.e., the program was taught between the first and second testing periods) or to the delayed treatment condition, (i.e., the program was taught between the second and third testing periods). Test responses generated at the first and second testing periods by the delayed treatment groups provided control data for the early treatment groups; during the delay period teachers continued with their regular geography program which included discussion of social issues without explicit attention to decision-making skills. After classrooms had been randomly assigned, one teacher requested a change from the late to early treatment condition for scheduling reasons unrelated to the pilot test. The resulting evaluation design is displayed in Table 3.

Students in each class were randomly assigned to three test order groups: one-third of the students in each class completed the tests in the order A-B-C, one-third in the order B-C-A and one-third in the order C-A-B.

A half-day inservice session was provided to the volunteer teachers to acquaint them with the growth schemes and the lesson plans. Instruction in marking the open-ended tests was provided at a subsequent inservice.

Student performance on each test occasion was machine scored for the multiple-choice booklets; (the maximum possible score was 26). The internal consistencies (Hoyt reliabilities) of the multiple-choice booklets on each test administration are displayed in Table 4.

The open-ended items were marked by the teachers, after inservice, by assigning a score that indicated the growth scheme level matched by the student response. (In a previous study [Ross & Maynes, 1983c] a similar procedure produced high

correlations between scores assigned by teachers and scores assigned to a random sample of responses by a trained marker.)

TABLE 3  
EVALUATION DESIGN

| Grade | <u>Early Treatment</u> |     | <u>Late Treatment</u> |     |
|-------|------------------------|-----|-----------------------|-----|
|       | Classes Student        |     | Classes Student       |     |
|       | n                      | n   | n                     | n   |
| 7     | 3                      | 82  | 3                     | 71  |
| 9     | 3                      | 62  | 2                     | 37  |
| 10    | 2                      | 51  | 1                     | 17  |
| Total |                        | 195 |                       | 125 |

TABLE 4  
INTERNAL CONSISTENCY (HOYT RELIABILITY) OF  
MULTIPLE-CHOICE INSTRUMENTS

| Test Occasion | Form A | Form B | Form C |
|---------------|--------|--------|--------|
| 1             | .44    | .62    | .58    |
| 2             | .57    | .71    | .68    |
| 3             | .64    | .70    | .61    |

Findings

The results displayed in Figure 1 and Table 5 suggest a very marked effect due to the program. A trend analysis of the 14 class means in the multiple-choice test indicated there was a strong linear trend in the data ( $F=52.78, p<.00001$ ) which was not different for the two groups ( $F=2.78, p<.12$ ). The quadratic trend, however, while nonsignificant ( $F=3.51, p<.10$ ) for the groups combined, was significantly different in the groups ( $F=43.36, p<.001$ ), reflecting the test occasion 2 differences in the means. There were significant differences in the overall test occasion scores ( $F=40.22, p<.00001$ ) and a significant interaction between test occasions and treatment groups ( $F=13.13, p<.001$ ) again reflecting test occasion 2, and possibly test occasion 3, group differences. The analysis also confirmed an overall difference between the two treatment groups ( $F=6.69, p<.05$ ).

Figure 1 displays these results and shows the large divergence in test scores on test occasion 2. The early treatment classes who received the instruction immediately following test occasion 1 show large gains on test occasion 2, while the

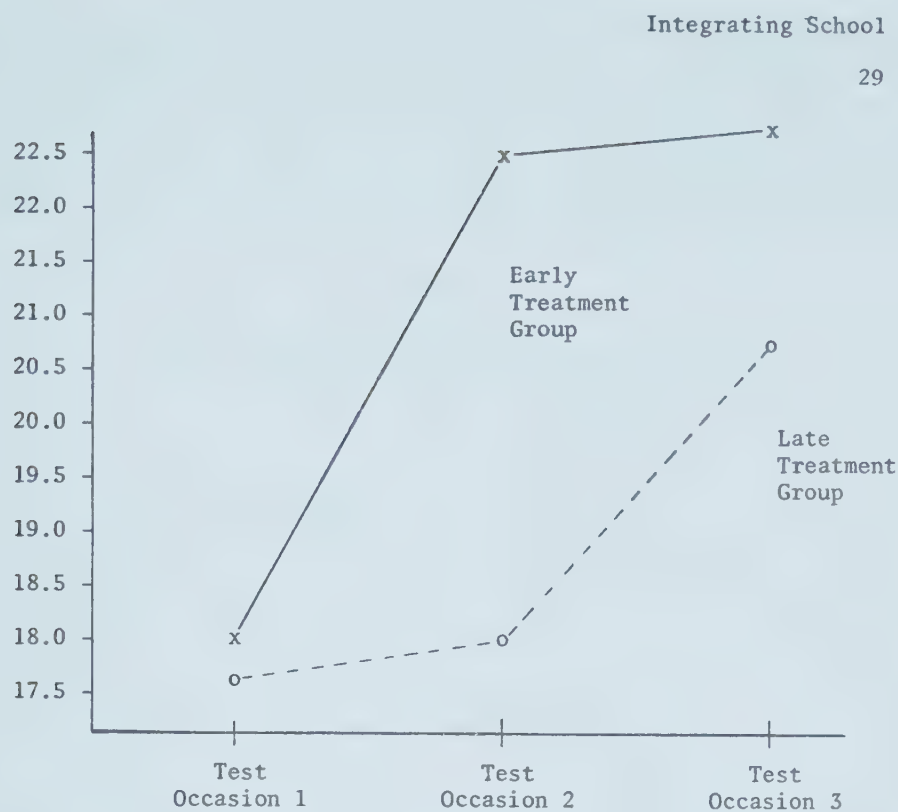


Figure 1. Overall Class Means on Multiple-Choice Instruments

uninstructed classes scored much as they did on pretest. By test occasion 3, the late treatment classes also show gains, and the early treatment classes show evidence of retention of what they had learned. The figure also shows that the late treatment classes, although performing very similarly to the early group on pretest, did not benefit as much from the instruction since their average test scores do not match those of the early group on test occasion 3. The list of class means in Table 5 shows that this is undoubtedly due to one anomalous grade 9 class in the late treatment group in which there was no change in the average score across all three occasions. Otherwise the direction of results is in every case, whatever the grade level, consistent with expectations.

The lower post-instructional performances of the late treatment students as compared to the early treatment group is reflected in the effect sizes (based on Glass, 1978, and computed using class means). Between test occasions 1 and 2 the early treatment elementary students showed an effect size between class means of 3.55, the secondary students of 4.49. In the late treatment, on the other hand, the effect sizes between occasions 2 and 3 were much smaller, 1.57 for the elementary students and 2.79 for the secondary students.<sup>1</sup>

The pattern of results across occasions for the open-ended items is very similar to that of the multiple-choice tests (see Figures 2 and 3). As would be expected for a produced rather than selected response, the scores tend to be lower. For the focusing items, on the open-ended pretest the students produced average responses of about level 1.5; on the multiple-choice tests they average around level 2. Post-instructional performance on the open-ended items ranges between level 2 and level 3 while on the multiple-choice items it tends to be level 4 and above. On the



frameworking tests there is a similar pattern with the open-ended pretests averaging less than 1.5 and increasing to around 2.5 on the posttest, while on the multiple-choice tests they moved from around level 3 to somewhat better than level 3.5. It was noted there was much greater variation in the subgroup performances on the multiple-choice framework items (results not displayed) than on the open-ended framework items, the latter showing the most consistent results of the study.

TABLE 5  
CLASS MEANS AND STANDARD DEVIATIONS  
MULTIPLE-CHOICE DATA

| Group           | Class | n  | Test Occasion |              |              |
|-----------------|-------|----|---------------|--------------|--------------|
|                 |       |    | 1             | 2            | 3            |
| Early Treatment |       |    |               |              |              |
| Elementary      |       |    |               |              |              |
| Grade 7         | 1     | 30 | 16.27 (3.61)  | 21.23 (3.71) | 21.33 (5.05) |
|                 | 2     | 30 | 16.30 (4.19)  | 21.87 (3.73) | 22.53 (3.38) |
|                 | 3     | 22 | 16.14 (3.43)  | 22.46 (3.33) | 22.23 (3.57) |
| Secondary       |       |    |               |              |              |
| Grade 9         | 4     | 17 | 15.82 (4.36)  | 20.53 (4.82) | 19.77 (3.82) |
|                 | 5     | 16 | 19.06 (3.47)  | 24.38 (3.38) | 24.38 (3.32) |
|                 | 6     | 29 | 18.55 (3.81)  | 21.45 (3.42) | 22.93 (2.99) |
| Grade 10        | 7     | 31 | 20.90 (3.73)  | 24.16 (3.31) | 23.90 (3.28) |
|                 | 8     | 20 | 21.60 (3.25)  | 23.70 (3.47) | 23.40 (3.86) |
| Late Treatment  |       |    |               |              |              |
| Elementary      |       |    |               |              |              |
| Grade 7         | 1     | 24 | 18.25 (3.95)  | 18.54 (4.11) | 21.79 (3.78) |
|                 | 2     | 19 | 15.47 (3.03)  | 15.95 (3.55) | 17.05 (4.67) |
|                 | 3     | 28 | 17.43 (3.28)  | 17.71 (4.01) | 23.61 (2.53) |
| Secondary       |       |    |               |              |              |
| Grade 9         | 4     | 22 | 19.18 (3.65)  | 19.82 (4.45) | 23.32 (4.12) |
|                 | 5     | 15 | 18.53 (3.96)  | 18.73 (4.22) | 18.53 (4.91) |
| Grade 10        | 6     | 17 | 17.88 (4.09)  | 17.24 (4.24) | 19.29 (4.58) |

Table 6 presents t-test comparisons for the two open-ended item subtests for each test order. The table shows that the comparisons between early and late treatment subgroups are as expected: there are no significant differences on test occasion 1; the early treatment subgroups outperform the late treatment subgroups on occasion 2; and with one exception there are no differences between subgroups on occasion 3. Similarly the table shows that the within-group comparisons are as expected: all the early treatment subgroups and none of the late treatment subgroups show significant gains on both skills from occasion 1 to occasion 2; with one exception all the early treatment subgroups show no change from occasion 2 to occasion 3, and with one exception all the late treatment subgroups show the expected gain from occasion 2 to occasion 3.

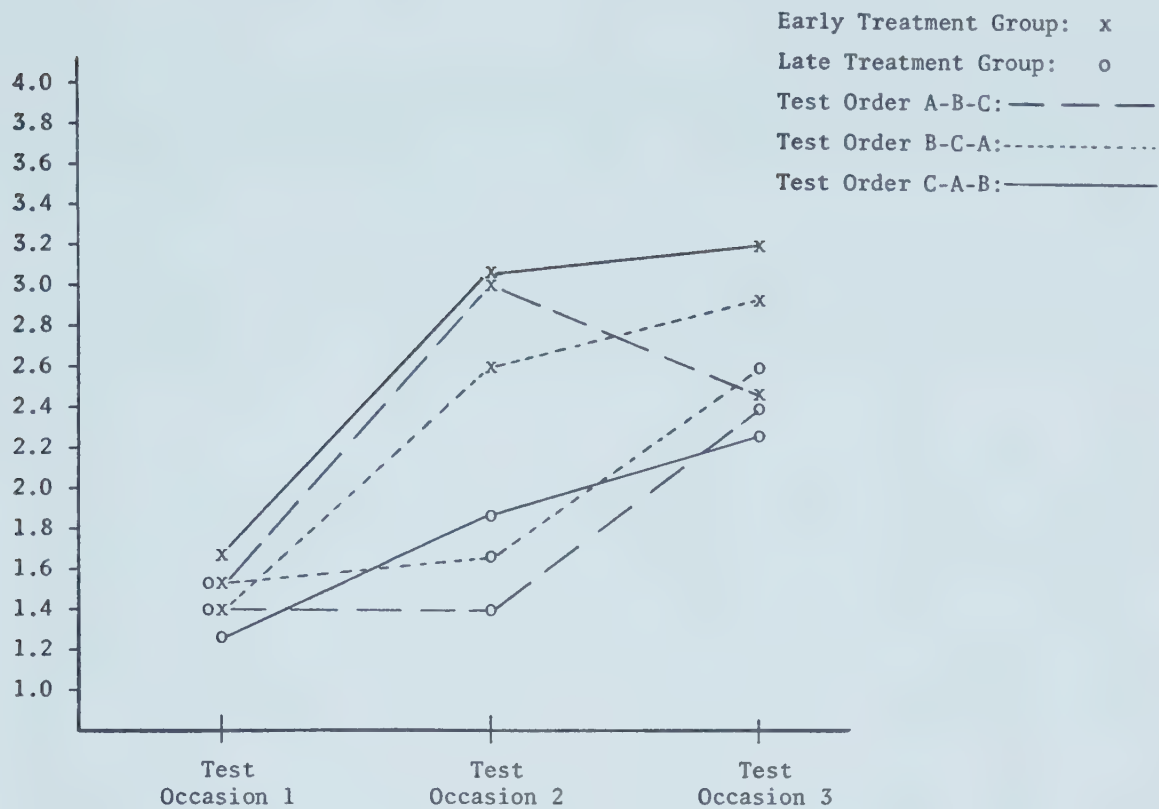


Figure 2. Open-Ended Focusing Test Results

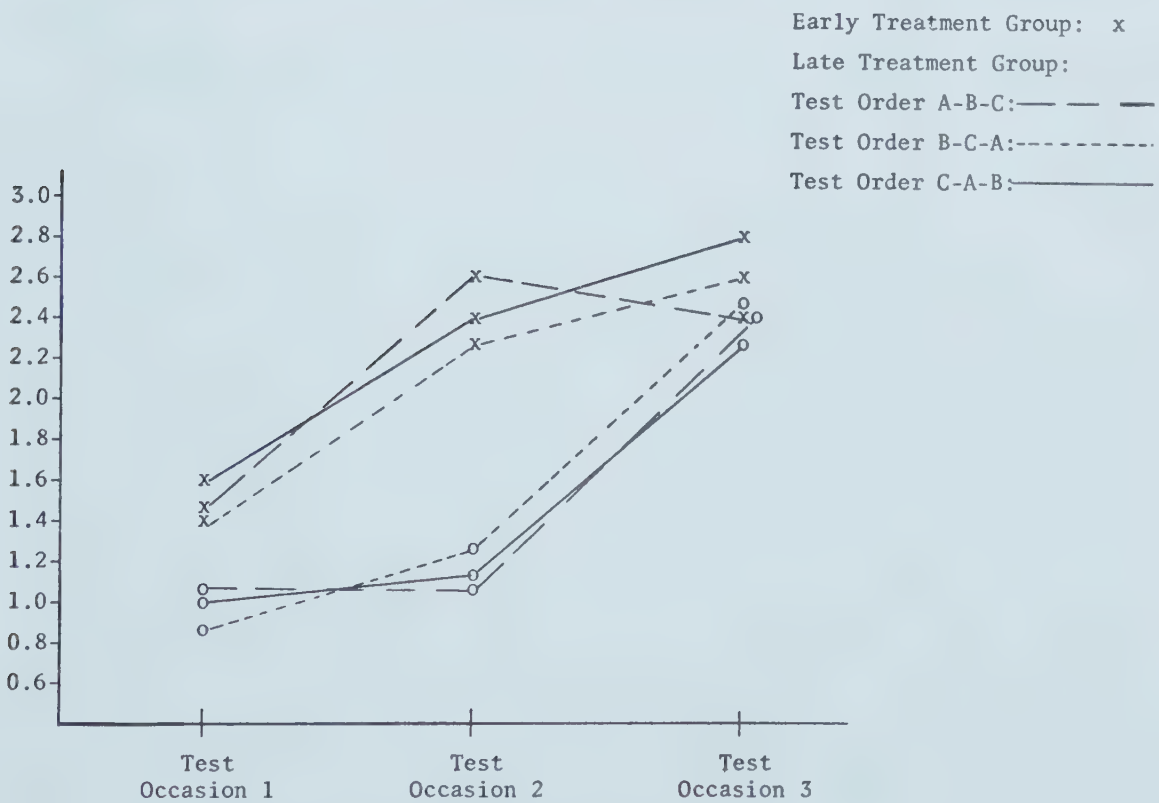


Figure 3. Open-Ended Frameworking Test Results

TABLE 6  
STUDENT MEANS AND STANDARD DEVIATIONS  
OPEN-ENDED SUBTESTS FOR EACH TREATMENT AND ORDER GROUP

| Order        | Group | n  | Test Occasions    |                    |                      |  |   |  |
|--------------|-------|----|-------------------|--------------------|----------------------|--|---|--|
|              |       |    | 1                 |                    | 2                    |  | 3 |  |
| Focussing    |       |    |                   |                    |                      |  |   |  |
| A-B-C        | early | 57 | 1.53 (1.54)<br>ns | 3.02 * (1.48)<br>* | 2.54 * (1.68)<br>ns  |  |   |  |
|              | late  | 39 | 1.39 (1.41)       | 1.44 ns (1.35)     | 2.41 * (1.55)        |  |   |  |
| B-C-A        | early | 57 | 1.37 (1.57)<br>ns | 2.63 * (1.63)<br>* | 2.93 ns (1.52)<br>ns |  |   |  |
|              | late  | 36 | 1.53 (1.58)       | 1.67 ns (1.39)     | 2.56 * (1.36)        |  |   |  |
| C-A-B        | early | 49 | 1.67 (1.61)<br>ns | 3.00 * (1.40)<br>* | 3.23 ns (1.23)<br>*  |  |   |  |
|              | late  | 41 | 1.29 (1.40)       | 1.85 ns (1.26)     | 2.27 ns (1.64)       |  |   |  |
| Frameworking |       |    |                   |                    |                      |  |   |  |
| A-B-C        | early | 57 | 1.51 (1.70)<br>ns | 2.60 * (1.95)<br>* | 2.39 ns (2.01)<br>ns |  |   |  |
|              | late  | 39 | 1.08 (1.58)       | 1.10 ns (1.48)     | 2.41 * (1.83)        |  |   |  |
| B-C-A        | early | 57 | 1.39 (1.47)<br>ns | 2.28 * (1.89)<br>* | 2.58 ns (1.82)<br>ns |  |   |  |
|              | late  | 36 | 0.89 (1.33)       | 1.33 ns (1.53)     | 2.47 * (1.80)        |  |   |  |
| C-A-B        | early | 49 | 1.57 (1.72)<br>ns | 2.37 * (2.00)<br>* | 2.80 ns (1.95)<br>ns |  |   |  |
|              | late  | 41 | 0.98 (1.73)       | 1.17 ns (1.50)     | 2.34 * (1.93)        |  |   |  |

\* Significant at  $p < .05$ , two tailed test of statistical significance

TABLE 7  
STUDENT MEANS AND STANDARD DEVIATIONS  
MULTIPLE-CHOICE AND OPEN-ENDED PRE- AND POSTTEST TESTS  
REPLICATION STUDY

| Test Order                  | n  | Test Occasion |                 |
|-----------------------------|----|---------------|-----------------|
|                             |    | 1             | 2               |
| Open-Ended Instruments      |    |               |                 |
| Focussing                   |    |               |                 |
| A-B                         | 83 | 1.81 (1.60)   | 2.62 * (1.63)   |
| B-A                         | 76 | 1.76 (1.49)   | 2.82 * (1.43)   |
| Frameworking                |    |               |                 |
| A-B                         | 83 | 1.33 (1.74)   | 2.57 * (1.80)   |
| B-A                         | 76 | 1.59 (1.63)   | 2.38 * (1.85)   |
| Multiple-Choice Instruments |    |               |                 |
| A-B                         | 83 | 18.51 (3.64)  | 23.23 * (4.06)  |
| B-A                         | 76 | 21.24 (3.57)  | 21.32 ns (3.59) |

\* Significant at  $p < .05$ , two-tailed test of statistical significance



### *Replication*

The pilot test was partially replicated in a second school system. Seven volunteer teachers taught the grade 7 version of the program. In this system no control group was formed and only two test forms, A & B, were used in a straight pre-posttest design. Half of the students in each of these classes received the tests in the A-B and half in the B-A order.

The results for the partial replication are displayed in Table 7 for both sets of instruments in both test orders. The posttest results show a statistically significant gain over pretest scores as anticipated with one exception (the B-A order for the multiple-choice scores).

### *Discussion*

The pilot test data provide sound evidence that the program was effective. After just a few hours of intensive instruction the responses of the students to problems requiring decisions have altered significantly, and moreover it appears that they retain what they have learned, at least for one month. This is true in all classes, regardless of grade or ability level, including, in the early treatment group, a grade 7 class that included a number of grade 6 students, and a grade 9 four-year class of lower ability students. The findings from the original study were partially replicated in a subsequent investigation involving a second school system. On the other hand, not all students show complete mastery of the skills. For that result to be achieved, teachers would have to discuss or assign many further example problems in other topics and contexts throughout the year.

To what extent does a cognitive approach reduce the previously described impediments to implementing the progressive ideal? There are several grounds for optimism.

1. Adopting a cognitive approach to the analysis of social issues broadens the group of potential users, making pursuit of the social concerns thrust more attractive to a larger proportion of the teacher population. Teachers with a modest interest in Social Reconstruction (using Eisner and Vallance's categories), but who give their main commitment to the development of cognitive processes, are the most likely beneficiaries of the broader appeal.

One should not exaggerate the extent to which the base of potential users could be expanded: curriculum orientations represent stable complexes of beliefs and ideals that govern curriculum choices of individual teachers. It is unlikely, for example, that those who were attracted to the anti-intellectualism of progressivism as manifested in the 1930s and '40s<sup>2</sup> would be moved by a cognitive appeal. By the same token, progressivism was an extremely pluralistic movement—one should not be excessively pessimistic about the possibilities of freeing the social concerns thrust from the excesses of the progressivists of this period.

2. The shortage of classroom resources is alleviated in part by the materials produced for the pilot test. These of course are written within an Ontario context and may not be directly transferable to other jurisdictions. So too, social issues materials quickly become dated as new information about important social concerns becomes available and the issues themselves rise and fall in public consciousness. The materials produced in the project are more likely to have utility indirectly as illustrations or as exemplary models of a strategy for teaching social issues analysis that can be rewritten to meet new curriculum contexts.

3. The main contribution of the cognitive approach is likely to be its surfacing of the underlying intellectual skills required to analyze social issues. By making these overt objectives of instruction and by providing students with explicit strategies appropriate to their levels of intellectual development, the likelihood of productive task oriented discussion increases substantially. Particularly relevant is the focusing of discussion that develops when students target their critiques of other positions taken in the class on specific cells of an alternatives x criteria matrix.

4. Finally, a cognitive approach provides greater specificity for judging the quality of the outcomes of social issues analysis. The specific criteria for making a decision (i.e., the dimensions of growth in Tables 1 and 2) provide an operational definition of the decision-making process which is a necessary (but not sufficient) condition of reflective social inquiry.

The study then has demonstrated that the skills can be taught, learned and used. No attempt was made to assess the nature and degree of concurrent content learning, nor did it assess the issue of the extent to which mastery of the skills assists students to integrate the knowledge accumulated in school and to recognize its usefulness in socially relevant contexts. But the expectation is plausible, the first steps have been accomplished, teachers are adopting the materials and further studies will monitor the outcomes.

#### Notes

1. In the absence of control group data for this group the effect size was calculated using the formula:

$$\frac{\bar{x} \text{ Test Occasion 3} - \bar{x} \text{ Test Occasion 2}}{\text{S.D. Test Occasion 2}}$$

2. The anti-intellectualism of progressivism in the 1930s and '40s is evidenced by a lack of interest in student performance on achievement measures (with the notable exception on the part of the testing movement led by Thorndike), by manipulating teachers and students through group dynamics to reach pre-determined ends in allegedly democratic discussions and pseudo-participatory decision-making, by misrepresenting the findings of pedagogical research and by subordinating disciplinary inquiry and classical knowledge structures to short term and trivial social needs of students. For a detailed account of this theme within progressivism, see Cremin (1963), Graham (1967, especially pp. 149-151), Nelson (1978), and Ravitch (1983, especially pp. 48-49 and pp. 68-80).

#### References

- Ausubel, D. P., & Robinson, F. G. (1969). *School learning: An introduction to educational psychology*. New York: Holt, Rinehart & Winston.
- Bourne, P., & Eisenberg, J. (1978). *Social issues in the curriculum*. Toronto: OISE Press.
- Broudy, H. S. (1977). Types of knowledge and purposes of education. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 1-17). Hillsdale, NJ: Erlbaum.
- Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology, Vol. 1* (pp. 77-165). New York: Wiley.
- Champagne, A. B., & Klopfer, L. E. (1981). Student beliefs about gravity and motion. *Problem-Solving*, 3(1), 12-13.
- Clement, J. (1981). Students' preconceptions in physics and Galileo's discussion of falling bodies. *Problem-Solving*, 3(1), 3-5.

- Cremin, L. A. (1962). *The transformation of the school: Progressivism in American education*. New York, NY: Alfred Knopf.
- Ebel, R. (1979). *Essentials of educational measurement* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Eisner, E. W., & Vallance, E. (Eds.). (1974). *Conflicting conceptions of curriculum*. Berkeley, CA: McCutchan.
- Glass, G. V. (1978). Integrating findings: The meta-analysis of research. In L. S. Shulman (Ed.), *Review of research in education*, Vol. 5 (pp. 351-379). Itasca, IL: Peacock.
- Graham, P. A. (1967). *Progressive education: From Arcady to academe. A history of the progressive education association 1919-1955*. New York, NY: Teachers College Press.
- Manson, G., & Vuicich, G. (1977). *Toward geographic literacy: Objectives for geographic education in the elementary school*. Boulder, CO: Social Science Education Consortium.
- Miller, J. P. (1983). *The educational spectrum*. New York: Longman.
- Nelson, M. R. (1978). Rugg on Rugg: His theories and his curriculum. *Curriculum Inquiry*, 8(2), 119-132.
- Newmann, F. A. (1975). *Education for citizen action: Challenge for secondary curriculum*. Berkeley, CA: McCutchan.
- Oliver, D. W., & Newmann, F. M. (1967-1973). *Public issues series*, 30 booklets. Columbus, OH: Xerox Educational Publications.
- Oliver, D. W., & Shaver, J. P. (1974). *Teaching public issues in the high school*. Logan, UT: Utah State University Press.
- Ravitch, D. (1983). *The troubled crusade: American education 1945-1980*. New York, NY: Basic Books.
- Roberts, D. A. (1982). Curriculum emphases—A key area for teacher decision-making (The case of school science). In K. A. Leithwood (Ed.), *Studies in curriculum decision-making* (pp. 226-244). Toronto: OISE Press.
- Ross, J. A. (1981). The measurement of student achievement in a decision-making approach to values education. *Alberta Journal of Educational Research*, 27(1), 1-15.
- Ross, J. A., & Maynes, F. J. (1981). *Geographic thinking skills*. Toronto: OISE Press.
- Ross, J. A., & Maynes, F. J. (1982). *Teaching problem-solving*. Toronto: OISE Press.
- Ross, J. A., & Maynes, F. J. (1983a). The development of a test of experimental problem-solving. *Journal of Research in Science Teaching*, 20(1), 63-75.
- Ross, J. A., & Maynes, F. J. (1983b). Teaching problem-solving: An instructional design strategy. *Canadian Journal of Education*, 8(2), 155-173.
- Ross, J. A., & Maynes, F. J. (1983c). Experimental problem-solving: An instructional improvement field experiment. *Journal of Research in Science Teaching*, 20(6), 543-556.
- Spiro, R. J. (1977). Remembering from text: The "state of schema" approach. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 137-165). Hillsdale, NJ: Erlbaum.



JOE BELANGER

*University of British Columbia*

and

R. GLENN MARTIN

*University of Alberta*

## The Influence of Improved Reading Skill On Writing Skill

*Four intact experimental classes (n=94) which received a reading treatment and four control classes (n=100) wrote pre- and post-compositions which were evaluated for overall quality, syntactic density, T-unit length, productivity, and selected grammatical errors. Despite substantial and statistically significant differences on the Davis Reading Test six months following the treatment, the post-compositions written by the two groups were remarkably similar, with random gains favoring first one group and then the other. Only the analysis of grammatical errors revealed changes consistently in the hypothesized direction, but most of these were statistically non-significant. Previous studies of the influence of reading skill on writing skill failed to demonstrate measurable reading changes on standardized instruments and therefore could claim no consequent writing changes. The current study achieved reading changes, but these did not result in changes on writing measures.*

### *Background*

One of the cornerstones of English pedagogy is the belief that reading and writing are interrelated, interdependent skills and the consequent assumption that that which enhances one will automatically benefit the other. Although this read-write pedagogy in its various forms can be traced back at least to the Sophists some two thousand years ago (Broudy & Palmer, 1965), empirical evidence to support a di-

---

Dr. Joe Belanger is an assistant professor of education at the University of British Columbia.

Dr. R. Glenn Martin is a professor of education at the University of Alberta. The research reported here is taken from Dr. Belanger's doctoral dissertation which was supported by a Canada Council Grant.

rect, causative relationship between reading skill and writing skill has remained elusive.

Despite considerable professional interest in the interrelationships among the language arts during the 1950s and 1960s (Hildreth, 1954; Artley, Hildreth, Townsend, Beery, & Dawson, 1954; Burrows, Parke, Edmund, Deboer, Horn, Herrick, & Strickland, 1961; Meckel, 1963; Russell & Fea, 1963; Harris, 1969), few researchers have conducted experimental studies into the relationships between reading and writing. Researchers attempting to teach reading skills and measure the effects on written composition have either failed to demonstrate significant improvement in reading skill (Eurich, 1931; Schneider, 1971) or failed to use standard measures of either reading or writing skill (Mathews, Larson, & Butler, 1945). On the other hand, studies by Obenchain (1971), Reed (1967), Nagle (1972), Weaver (1977), Levine (1976), Combs (1977), Collins (1981), and Yusuf (1982) suggest that various writing programs may influence reading skill.

Tangential evidence on the relationships between the skills is provided by descriptive studies (Loban 1963, 1966, and 1967; Diederich, 1957; Lazdowski, 1976; Baker, 1954; Evanechko, Ollila, & Armstrong, 1974; Heller, 1980; Shanahan, 1980; McConnell, 1982), by correlational studies (Kuntz, 1975; Johnson, 1976; Harris, 1975; Fishco, 1966; Schonell, 1942; Thomas, 1976; Hill, 1982; Mohler, 1983), and by studies of the effects of various methods of teaching beginning reading (Fyfe, 1965; Shapiro & Shapiro, 1973; Mazurkiewicz, 1973; Stewart, 1969; Ackerman, 1969; Smith, 1968; Quinn, 1977; Eckhoff, 1983). The most careful documentation of this relationship is provided by Loban's 13-year study in the Oakland Public Schools. Commenting on the reading and writing abilities of grade-six students, Loban (1963) noted that "on every statistical measure, one fact is extremely clear in the present study: those who read well also write well; those who read poorly also write poorly" (p. 75). Reporting on the same students when they had reached grade nine, he noted that the relationships became more pronounced as the years passed. Diederich (1957), examining the relationship between verbal scores on the Scholastic Aptitude Test and teachers' quality ratings of students' composition skills, concluded that good measures of reading ability were the most trustworthy indicators of writing ability. A good deal of correlational data also supports the hypothesis that reading and writing are closely related skills. Tovatt and Miller (1967) reported very high correlations among the STEP Reading, Writing, and Listening tests with coefficients ranging from .86 to .91. Correlation coefficients between these tests and written compositions ranged from .44 to .62. Theorists such as Keltner (1957), Becker (1963), and Thompson (1967), on the other hand, argue that such correlations are in fact not meaningful and that they may well be taken to indicate differences rather than similarities among the skills. In addition, a number of studies have failed to show relationships between reading and writing skills (Siedow, 1973; Fuller, 1974; DeLuca, 1979; Perry, 1980; MacNeill, 1982, for example). Stotsky's (1983) review of the literature shows little progress on the basic issues involved since the survey of studies by one of the present authors five years earlier (Belanger, 1978b). At the present state of knowledge, the striking correlations between reading and writing remain poorly understood, despite their apparent promise in the troubled curricular domain of literacy. Hence the urgency to go beyond correlations to any level of causal interrelationships that may be involved.

### *The Study*

This study was designed to assess the possibility that reading change can produce writing change. It examined the relationship between measured changes in reading ability and consequent measured changes in writing skill, if any, at two grade levels over a six-month period. The design called for two experimental and two control classes at each of the grade-nine and grade-ten levels. Following the administration of the initial reading test and collection of the initial writing samples, the experimental students received the reading treatment. Parallel reading tests were given and parallel writing samples were taken three months and six months following the administration of the treatment.

### *Hypotheses*

We hypothesized that groups of experimental students who had shown significantly greater gains on reading measures than control groups did would, as a consequence, show greater gains on writing measures than the control groups did. We asked specifically:

1. Would the hypothesized writing gains be most evident in measures of quality, syntactic maturity (T-unit length or syntactic density), productivity, or skill in usage?
2. Would the hypothesized writing gains follow the reading gains immediately (be most evident in the first three months)? follow the reading gains but with some delay (be most evident in the second three months)? or accompany the reading gains (be evident at both the three- and six-month periods)?
3. Would the hypothesized writing gains be influenced by the sex of the subject, the ability of the subject, or the grade level of the subject?

### *Procedures*

The experimental and control groups were compared on the reading and each of the four writing measures using three time periods: the initial three months (the pre-mid comparison), the second three months (the mid-post comparison), and the full six months of the experiment (the pre-post comparison). Since pre-mid and pre-post results were very similar, only the six-month comparison (pre-post) is presented here. Data on the 36 null hypotheses tested in the study—comparisons on the four writing measures by grade level, treatment, and sex of the subject for each of the three time periods—are available in the complete document (ERIC: ED 163 409).

1. *The Sample.* Eight intact classes, two pairs on the grade-nine and two pairs on the grade-ten level, each pair taught by one teacher, were randomly designated as experimental and control groups. After attrition, 94 experimental and 100 control students were involved from November through June.

2. *The Treatment.* This study was conducted in conjunction with a reading experiment that examined longitudinal effects of the S.O.S. Reading Technique, a replication on a larger scale of several previous studies (Martin, 1975). The reading study was conducted and reported by Martin and Belanger (1977). The writing study differs from previous read-write studies, in that differences in reading growth between the experimental and control groups were assured prior to undertaking analysis of the writing sample. Although the data for the reading and writing studies



were collected simultaneously, writing analysis was delayed until reading results were clear. A detailed rationale for the reading technique is contained in E. Martin (1976) and R. Martin (1977).

Three forms of the *Davis Reading Test*, Series 2, were used as the reading measure. Reviews of the test (Buros, 1965) suggested that it was the best instrument with three or more forms available for the grade levels in question.

TABLE 1  
READING MEASURE: STATISTICAL DECISIONS BASED ON ANALYSIS OF COVARIANCE FOLLOWED BY PRE-POST CHANGES ON THE READING MEASURE

| Statistical<br>Data    | Grade 9 |           | Grade 10 |           | Grades Combined |           |
|------------------------|---------|-----------|----------|-----------|-----------------|-----------|
|                        | n       | prob      | n        | prob      | n               | prob      |
| Treatment              | 103     | 0.020*    | 91       | 0.023*    | 194             | 0.0013*   |
| Teacher<br>Interaction | 103     | 0.16      | 91       | 0.69      | 194             | 0.55      |
| Sex<br>Interaction     | 103     | 0.93      | 91       | 0.48      | 194             | 0.65      |
| Average Change         |         |           |          |           |                 |           |
| Experimental           | n       | $\bar{x}$ | n        | $\bar{x}$ | n               | $\bar{x}$ |
| Boys                   | 28      | +0.82     | 25       | +3.60     | 53              | +2.13     |
| Girls                  | 22      | +3.46     | 19       | +1.00     | 41              | +2.32     |
| Combined               | 50      | +1.98     | 44       | +2.48     | 94              | +2.21     |
| Control                |         |           |          |           |                 |           |
| Boys                   | 24      | +0.13     | 23       | -1.61     | 47              | -0.72     |
| Girls                  | 29      | -1.03     | 24       | -0.25     | 53              | -0.68     |
| Combined               | 53      | -0.51     | 47       | -0.92     | 100             | -0.70     |

\*p < .05

3. *Findings of the Reading Study.* As Table 1 shows, with the pretest as covariate, experimental and control groups are significantly different on the posttest, both when the grades are examined separately and when they are combined. Over the six months of the experiment, growth in reading skill was divided equally between the grade-nine and grade-ten subjects. Nor was sex a significant factor in the results, as lack of significant interaction between the treatment and the sex of the subjects shows. The three-month comparison (not shown) revealed that the majority of these gains occurred in the first three months of the experiment although the gap between the experimental and control groups continued to widen, albeit more slowly, over the second three months.

In addition, to test for differing effects with high-, medium-, and low-ability students, two way analysis of variance was used. No statistically significant interaction between treatment and ability level was found. Probabilities ranged between .40

and .70 suggesting that the effects of the treatment were not influenced by the subjects' ability.

The writing study depended, of course, on the nature of the reading results. These were statistically significant, but were they large enough to be able to generate consequent writing changes? One way of examining the magnitude of the reading change is to compare the data with between-year differences on the norms which accompany the reading test. As Table 1 shows, over the six months of the experiment, the grade-nine and -ten experimental students gained 2.49 and 3.40 points respectively more than their control counterparts. A "year's" difference on the Davis norms is a rough approximation and varies with the grades involved and the levels of the raw scores. As a general rule, about six raw-score points separate grade-nine and -ten levels on the raw scores. By this measure, the experimental students exceeded the control students by about a half-year's growth, .41 for the grade nines and .57 for the grade tens.

4. *The Writing Measure.* One 30-minute in-class composition was written on assigned topics at each of the three test periods. Writing topics were adapted from Bell's (1971) study of the writing preferences of high school students in the Edmonton area, a study showing that grade-ten students voluntarily chose "school," "home," and "society" most often as essay topics. The three topics were systematically rotated at each testing period.

Student papers were subjected to five separate analyses: a 37-percent random subsample of the compositions was judged for overall quality by experienced raters; this subsample was reanalyzed for major sentence errors; an additional 13 percent subsample (that is a total of 50 percent) was analyzed using Belanger's (1978a) correction of Golub and Kidder's (1974) Syntactic Density Score (SDS); and all of the papers collected were analyzed for T-unit length and productivity.<sup>1</sup>

TABLE 2  
RATER RELIABILITIES OBTAINED EACH DAY USING  
PEARSON PRODUCT-MOMENT CORRELATIONS

| Day      | Three Raters |
|----------|--------------|
| 1        | .868         |
| 2        | .881         |
| Grade 10 | .868         |
| 3        | .847         |
| 4        | .876         |
| Grade 9  | .859         |
| Overall  | .859         |

a. *Overall Quality Analysis.* The papers of 36 experimental and 36 control students (216 compositions) were judged for overall quality by three experienced raters using Diederich's (1974) scoring procedures under conditions outlined by Braddock,

Lloyd-Jones, and Schoer (1963). As Table 2 shows, Diederich's (1974) three-rater method resulted in an overall interrater reliability of .858 calculated by Pearson product-moment correlations. Calculations were based on the two ratings ultimately used to establish the grade for each paper.

b. *Infrequency of Major Errors.* The infrequency of major usage errors (fragments, run-on sentences, errors in subject-verb agreement, and errors in pronoun agreement) was found to be a significant predictor of "good" tenth-grade writing by Martin (1968) and to correlate highly with teachers' quality ratings by Harris (1975) and Grobe (1981). Though significant changes in this area were not hypothesized, a subsample of the papers (the 216 compositions used in the quality analysis above) was analyzed for major errors.

c. *Syntactic Density Score.* The papers of 12 experimental and 12 control students were added to those in the overall quality subsample for the Syntactic Density Score analysis, a total of 288 compositions. Golub and Kidder's (1974) Syntactic Density Score has T-unit length as a basic component but goes beyond clause and sentence length factors and includes such items as number of modals, number of "be" and "have" auxiliaries, the number of prepositional phrases, the number of possessives, the number of adverbs of time, and the number of gerunds, participles and absolutes found in structures of free modification. Belanger's (1978a) correction eliminates a mathematical anomaly whereby an increase in the number of T-units analyzed decreased the score of a writing sample. A check-coder analyzed photocopies of 20 randomly-selected compositions.

d. *T-unit Length.* All 582 compositions gathered in the study were analyzed for Hunt's (1965) T-unit length, said by O'Donnell (1976) to be a simple, objective, and valid measure of syntactic maturity. Since T-unit length is a subscore of the SDS, the check-coder's ratings were abstracted from the check-coding for the SDS and compared with the investigator's ratings. The two assessors disagreed on the number of T-units in only one of the 20 compositions.

e. *Productivity.* Martin (1968), Miller and Ney (1968), Zanotti (1970), Obenchain (1971), Palmer (1971), Stewart and Grobe (1979), and Grobe (1981) report that the length of a time-controlled composition is a good indicator of its quality. Again, check-codings were abstracted from the SDS.

5. *Correlations Among Writing Measures.* To estimate the degree of similarity among the writing measures, Pearson product-moment correlations were calculated. As Table 3 shows, the quality and productivity measures correlate very highly with each other and the syntactic density and T-unit measures correlate very highly with each other. However, the quality/productivity and syntactic density/T-unit measures correlate only slightly—often negatively—with each other. The very high correlations of beyond .90 between T-unit length and the SDS support findings of O'Donnell (1976) who suggests that the measures are so closely related (because the SDS relies so heavily on T-unit length) that the extra work required to calculate the SDS does not appear to be justified. Even with the corrected scale used here but not by O'Donnell, the two measures appear to yield virtually the same kinds of results. The slight negative correlations between the quality and T-unit measures support findings by Potter (1966), Martin (1968), Stewart and Grobe (1979), and Crowhurst (1980) which suggest that T-unit length is not a good indicator of composition quality.

6. *Statistical Procedures.* Because the initial reading measure indicated that the experimental and control groups were not of equivalent ability at the beginning of



the experiment, analysis of covariance with the pretest as the covariate was used to compare the experimental and control groups on both the reading and writing measures. Interactions between treatment and teacher and between the treatment and the sex of the subject were examined. In addition, two-way analysis of variance was used to examine the effects of the treatment on high-, mid-, and low-ability students, and correlations were computed between the reading and writing measures. The analysis of variance and covariance were performed at the University of Alberta Computing Centre using the program "ANOV 35: Three-Way Analysis of Variance and Covariance." Pearson product-moment correlations were computed on the program "DESTO5: Correlations with Optional t-tests."

TABLE 3  
CORRELATIONS AMONG THE FOUR WRITING MEASURES USED IN THE STUDY

|                   |      | Quality |        |        | Syntactic Density |        |       | T-Unit Length |        |       |
|-------------------|------|---------|--------|--------|-------------------|--------|-------|---------------|--------|-------|
|                   |      | n       | r      | prob   | n                 | r      | prob  | n             | r      | prob  |
| Syntactic Density | Pre  | 72      | -0.069 | 0.562  |                   |        |       |               |        |       |
|                   | Mid  | 72      | -0.029 | 0.809  |                   |        |       |               |        |       |
|                   | Post | 72      | -0.020 | 0.865  |                   |        |       |               |        |       |
| T-Unit Length     | Pre  | 72      | -0.072 | 0.547  | 96                | 0.944  | 0.001 |               |        |       |
|                   | Mid  | 72      | -0.059 | 0.623  | 96                | 0.901  | 0.001 |               |        |       |
|                   | Post | 72      | -0.031 | 0.797  | 96                | 0.904  | 0.001 |               |        |       |
| Fluency           | Pre  | 72      | 0.347  | 0.003  | 96                | -0.068 | 0.508 | 194           | 0.009  | 0.903 |
|                   | Mid  | 72      | 0.483  | 0.0001 | 96                | -0.255 | 0.012 | 194           | -0.174 | 0.015 |
|                   | Post | 72      | 0.425  | 0.0001 | 96                | 0.087  | 0.399 | 194           | 0.094  | 0.190 |

<sup>a</sup> Since the sample sizes vary from 72 to 194, both the correlation coefficients (r) and the probabilities are given.

Findings

The data presented in this section offer absolutely no support for the hypothesis that growth in reading skill will automatically be accompanied by growth in writing skill. Of the 36 null hypotheses tested in the study, only three were rejected at the .05 level of confidence, approximately the number of rejections that would be expected by chance alone. Nor were these differences in the hypothesized direction: all three of the significant differences favored the control group, but the non-significant differences were almost randomly distributed between the two groups, some comparisons favoring the experimental group, some favoring the control group, with no apparent pattern of changes emerging. A subsidiary analysis of major errors found in the compositions was the only analysis to show promise: although only one of the nine comparisons showed statistically significant differences, all comparisons favored the experimental group. On the other hand, significant positive correlations between the reading and both the quality and productivity measures (but not the T-unit or syntactic density measures) suggest that some common basis for the two skills exists, confirming findings of a number of previous studies.<sup>2</sup> Statistically significant interactions between the treatment and the sex of the subject on 5 of the 36 comparisons (about 14 percent) suggest that the sex of the subject might have had some minor influence on the results.

TABLE 4

QUALITY MEASURE: STATISTICAL DECISIONS FOLLOWED BY  
MEAN PRE-POST CHANGES BY GRADE AND TREATMENT

| <u>Statistical Data</u> |    |    |         |             |  |          |
|-------------------------|----|----|---------|-------------|--|----------|
|                         | n  | df | F-ratio | Probability |  | Decision |
| Grade 9                 | 36 | 1  | 0.593   | .45         |  | N.S.     |
| Grade 10                | 36 | 1  | 1.658   | .21         |  | N.S.     |

| <u>Mean Changes</u> |         |           |          |           |          |                              |
|---------------------|---------|-----------|----------|-----------|----------|------------------------------|
|                     | Pretest |           |          | Posttest  |          | Posttest<br>Minus<br>Pretest |
|                     | n       | $\bar{x}$ | $\sigma$ | $\bar{x}$ | $\sigma$ |                              |
| Grade 9 Exp         | 18      | 5.64      | 1.13     | 5.42      | 1.29     | -0.22                        |
| Grade 9 Con         | 18      | 5.41      | 1.88     | 5.66      | 1.80     | +0.25                        |
| Grade 10 Exp        | 18      | 5.90      | 1.34     | 5.31      | 1.60     | -0.59                        |
| Grade 10 Con        | 18      | 6.16      | 1.31     | 5.79      | 1.64     | -0.37                        |

Table 4 presents the six-month comparison on the writing quality measure for the grade-nine and grade-ten groups separately. Inspection of Table 4 reveals no statistically significant differences between the experimental and control groups on either the grade-nine or grade-ten levels over the six months of the experiment. Examination of the data shows that small changes in quality scores favored the control groups on both grade levels. These data simply continue a trend observed on the three-month comparison: small, positive changes on the measure favored the control groups. On the grade-ten level there was a significant ( $p<.02$ ) interaction between the treatment and the sex of the subject, an interaction attributed to compositions written by three control girls who scored below the mean on the pretest but above the mean on both the mid- and posttests. This must be regarded as spurious in light of the large number of interactions tested. There was no statistically significant interaction between the treatment and the ability of the subject.<sup>3</sup>

Table 5 reveals no statistically significant differences between the experimental and control groups on the syntactic density measure on either the grade-nine or grade-ten levels. Over the six months of the experiment, small changes in the syntactic density scores favored the experimental groups. These data are at variance with trends observed in the first three months of the experiment where equally small and statistically non-significant differences favored the control groups. The grade-level conversion scale accompanying the SDS shows that a year's growth is approximately .8 points, suggesting that the differences reported above of .2 points maximum are quite minor. Nor did any pattern emerge on the pre-mid, mid-post, or pre-post comparisons to suggest that delayed improvement on the SDS measure might be expected.

TABLE 5

SYNTACTIC DENSITY MEASURE: STATISTICAL DECISIONS FOLLOWED BY MEAN PRE-POST CHANGES BY GRADE AND TREATMENT

| <u>Statistical Data</u> |    |    |         |             |          |
|-------------------------|----|----|---------|-------------|----------|
|                         | n  | df | F-ratio | Probability | Decision |
| Grade 9                 | 48 | 1  | 0.704   | .41         | N.S.     |
| Grade 10                | 48 | 1  | 0.672   | .42         | N.S.     |

| <u>Mean Changes</u> |                |           |          |                 |          |                              |
|---------------------|----------------|-----------|----------|-----------------|----------|------------------------------|
|                     | <u>Pretest</u> |           |          | <u>Posttest</u> |          | Posttest<br>Minus<br>Pretest |
|                     | n              | $\bar{x}$ | $\sigma$ | $\bar{x}$       | $\sigma$ |                              |
| Gr. 9 Exp           | 24             | 3.51      | .67      | 3.54            | .74      | +.03                         |
| Gr. 9 Con           | 24             | 3.70      | .74      | 3.50            | .96      | -.20                         |
| Gr. 10 Exp          | 24             | 3.64      | .90      | 3.87            | .79      | +.23                         |
| Gr. 10 Con          | 24             | 3.78      | 1.25     | 3.88            | .87      | +.10                         |

TABLE 6

T-UNIT MEASURE: STATISTICAL DECISIONS FOLLOWED BY MEAN PRE-POST CHANGES BY GRADE AND TREATMENT

| <u>Statistical Data</u> |     |    |         |             |          |
|-------------------------|-----|----|---------|-------------|----------|
|                         | n   | df | F-ratio | Probability | Decision |
| Gr. 9.                  | 103 | 1  | .286    | .59         | N.S.     |
| Gr. 10                  | 91  | 1  | 2.202   | .14         | N.S.     |

| <u>Mean Changes</u> |                |           |      |                 |      |                              |
|---------------------|----------------|-----------|------|-----------------|------|------------------------------|
|                     | <u>Pretest</u> |           |      | <u>Posttest</u> |      | Posttest<br>Minus<br>Pretest |
|                     | n              | $\bar{x}$ |      | $\bar{x}$       |      |                              |
| Gr. 9 Exp           | 50             | 13.97     | 2.84 | 14.45           | 2.78 | +.48                         |
| Gr. 9 Con           | 53             | 14.51     | 2.70 | 14.81           | 3.09 | +.32                         |
| Gr. 10 Exp          | 44             | 15.62     | 3.11 | 14.99           | 3.31 | -.63                         |
| Gr. 10 Con          | 47             | 14.87     | 3.24 | 14.91           | 3.37 | +.04                         |



Table 6 reveals no statistically significant differences between the experimental and control groups on the T-unit measure over the six months of the experiment. On the grade-nine level, small differences favored the experimental group while on the grade-ten level, somewhat larger but statistically non-significant differences favored the control groups. The comparison between the two groups over the first three months showed a significant treatment effect: The grade-ten control group was significantly better ( $p<.04$ ) than the experimental group as a result of “losses” experienced by the experimental group. Such “losses” might well result from the fact that the experimental group wrote T-units almost one full word longer than those written by the control group on the pretest. There was also a significant ( $p<.02$ ) interaction between the treatment and the sex of the subject on both the first-three-month and full-six-month comparisons, an interaction attributed largely to the work of three grade-ten control girls whose papers improved dramatically following the pretest.

TABLE 7  
PRODUCTIVITY MEASURE: STATISTICAL DECISIONS FOLLOWED BY  
MEAN PRE-POST CHANGES BY GRADE AND TREATMENT

| Statistical Data |     |    |         |      |          |  |
|------------------|-----|----|---------|------|----------|--|
|                  | n   | df | F-ratio | Prob | Decision |  |
| Gr. 9            | 103 | 1  | .668    | .42  | N.S.     |  |
| Gr. 10           | 91  | 1  | .126    | .72  | N.S.     |  |

| Mean Changes | Pretest |           |          | Posttest  |          | Posttest<br>minus<br>Pretest |
|--------------|---------|-----------|----------|-----------|----------|------------------------------|
|              | n       | $\bar{x}$ | $\sigma$ | $\bar{x}$ | $\sigma$ |                              |
| Gr. 9 Exp    | 50      | 245.4     | 96.5     | 213.1     | 90.4     | -32.3                        |
| Gr. 9 Con    | 53      | 266.9     | 80.3     | 231.0     | 68.5     | -35.9                        |
| Gr. 10 Exp   | 43      | 280.1     | 102.6    | 256.5     | 85.2     | -23.6                        |
| Gr. 10 Con   | 47      | 274.6     | 113.2    | 256.9     | 114.8    | -17.7                        |

Table 7 presents the six-month comparison on the productivity measure. There were no significant differences at either grade level in the length of the compositions written by the experimental and control groups over the six months of the experiment. Inspection of the raw scores shows that slight differences favored the experimental group on the grade-nine level and the control group on the grade-ten level. On the first three-month comparison, substantial but statistically non-significant differences averaging 27 words per composition favored the experimental groups on both the grade-nine and grade-ten levels, differences which evaporated by the end of six months. All groups wrote shorter posttests than pretests, a phenomenon, as noted above, which may be attributable to the attitude of the students in June.

As Table 8 shows, there were no significant differences between experimental and control groups in reducing major errors at either grade level. However, in both grades, differences favored the experimental groups. Interestingly, very slight dif-

ferences over the first three months favored the control groups on both grade levels while much more substantial (albeit statistically non-significant) differences over the final three months favored the experimental groups. It should also be noted that significant differences were not hypothesized here because these conventions of standard usage were assumed to be taught or induced over a long period of time. While no further inference is warranted, these differences between the experimental and control groups come closer than any others in the study to showing the hypothesized but uncorroborated influence of reading skill on writing skill.

TABLE 8  
ERROR REDUCTION: t-TEST RESULTS FOLLOWED BY  
MEAN PRE-POST CHANGES BY GRADE AND TREATMENT

| <u>t-tests</u> |    | Experimental  |  | t      | Probability | Decision |
|----------------|----|---------------|--|--------|-------------|----------|
|                | df | Minus Control |  |        |             |          |
| Gr. 9          | 34 | -0.34         |  | -1.024 | .40         | N.S.     |
| Gr. 10         | 34 | -0.60         |  | -1.677 | .20         | N.S.     |

| <u>Mean Changes</u> |    | Pretest   |          | Posttest  |          | Posttest<br>minus Pretest |
|---------------------|----|-----------|----------|-----------|----------|---------------------------|
|                     | n  | $\bar{x}$ | $\sigma$ | $\bar{x}$ | $\sigma$ |                           |
| Gr. 9 Exp           | 18 | 0.95      | 0.59     | 0.60      | 0.64     | -0.35                     |
| Gr. 9 Con           | 18 | 1.19      | 0.88     | 1.18      | 0.77     | -0.01                     |
| Gr. 10 Exp          | 18 | 1.45      | 1.40     | 0.99      | 1.06     | -0.46                     |
| Gr. 10 Con          | 18 | 0.83      | 0.81     | 0.97      | 1.19     | +0.14                     |

TABLE 9  
CORRELATIONS BETWEEN THE READING AND THE FOUR WRITING MEASURES  
AT THE THREE TEST PERIODS WITH SEPARATE CALCULATIONS FOR  
THE BOYS AND THE GIRLS ON THE QUALITY AND PRODUCTIVITY MEASURES

| WRITING<br>MEASURES     | READING MEASURE |         |       |         |       |          |       |
|-------------------------|-----------------|---------|-------|---------|-------|----------|-------|
|                         | n               | Pretest |       | Midtest |       | Posttest |       |
|                         |                 | r       | prob  | r       | prob  | r        | prob  |
| Quality                 | 72              | .47     | .0001 | .54     | .0001 | .48      | .0001 |
| Syntactic<br>Density    | 96              | .02     | .82   | .07     | .51   | .09      | .34   |
| T-unit length           | 194             | .06     | .43   | .08     | .28   | .09      | .19   |
| Productivity            | 194             | .20     | .004  | .19     | .008  | .27      | .0001 |
| Quality (Boys)          | 33              | .35     | .05   | .33     | .06   | .37      | .03   |
| Quality (Girls)         | 39              | .57     | .0001 | .70     | .0001 | .59      | .0001 |
| Productivity<br>(Boys)  | 100             | .26     | .01   | .15     | .13   | .36      | .0001 |
| Productivity<br>(Girls) | 94              | .15     | .14   | .22     | .03   | .17      | .10   |

As Table 9 shows, the highest correlations observed for the total sample were between the reading and quality measures. These moderately high correlation coefficients, ranging from .47 to .54 were significant beyond the .0001 level of confidence. The productivity measure also correlated significantly ( $p<.008$ ) with the reading measure, but the correlation coefficients were low (.19 to .27). The correlations between the reading and the syntactic density or T-unit measures, on the other hand, were all very small, none with a correlation coefficient reaching .10.

Table 9 also shows that correlations between reading and quality measures were much higher for the girls in the study than for the boys. Correlation coefficients for the girls ranged from .57 to .70 ( $p<.0001$ ) while those for boys ranged from .33 ( $p>.05$ ) to .37 ( $p<.05$ ). This supports earlier data reported by Fishco (1966) who found that correlations between reading and writing measures for the girls in his study were significant beyond the .01 level of confidence, but those for the boys did not reach the .05 level of confidence. On the other hand, on the productivity measure, the correlations for the boys were generally somewhat higher than those for the girls. This may well result from a small number of poor male writers who wrote very short papers.

TABLE 10  
RANKINGS OF THE GRADE 9 AND 10 EXPERIMENTAL AND CONTROL BOYS  
AND GIRLS ON THE INITIAL READING AND WRITING QUALITY MEASURES

|     |       | Grade 9                    |                 | Grade 10                   |                            |
|-----|-------|----------------------------|-----------------|----------------------------|----------------------------|
|     |       | Writing<br>Quality<br>Rank | Reading<br>Rank | Writing<br>Quality<br>Rank | Reading<br>Quality<br>Rank |
| Exp | Boys  | 1                          | 4               | 4                          | 3                          |
| Exp | Girls | 4                          | 3               | 1                          | 4                          |
| Con | Boys  | 2                          | 1               | 2                          | 2                          |
| Con | Girls | 3                          | 2               | 3                          | 1                          |

As Table 10 shows, when the grade-nine and grade-ten experimental and control boys and girls are ranked on the pre-reading and pre-writing quality measures, a group's rank on one measure can be used to predict its rank on the other measure in only one of the eight possibilities.<sup>4</sup> While in some cases there were only small differences in actual scores between two rankings, in other cases the differences were substantial. The grade-ten experimental girls, for example, scored almost one standard deviation below the control girls on the initial reading measure, yet one standard deviation above the control girls on the initial writing quality measure. Nor does this appear to be an anomaly of testing as the experimental girls maintained the high-quality and low-reading scores on both the mid- and posttests. The grade-nine experimental boys showed a similar profile, scoring one-half standard deviation below the grade-nine mean on the initial reading measure and one-half standard deviation above the mean on the writing measure, also maintaining this profile on the mid- and posttests. Comparative rankings for the reading and the other three writing measures are substantially the same as those reported in Table 10. Although such group rankings are crude comparisons, they suggest that the moderately high correlations ranging from .47 to .70 reported above may be



misleading. The correlations imply that skill in one area will be accompanied by skill in the other. The rankings, on the other hand, show that when class-sized groups are examined, the best writers are not necessarily the best readers—indeed, as a group, they seldom are. These findings corroborate Loban's (1963) observation that many students were good at one skill but poor at the other. For example, in Loban's study one-quarter of the highest fourth-grade readers were inferior or illiterate writers while one-third of the poorest writers were good or superior readers. He found similar patterns at other grade levels.

### Discussion

This study could find no evidence that reading skill has a direct influence on writing skill. Statistically significant and substantial improvement on the reading measure, improvement which was spread evenly among the grade nines and grade tens and the boys and girls in the study, improvement which was most evident in the first three months of the study and which continued to increase slightly over the second three months, appeared to exert no influence on the compositions written by the students when these compositions were examined for overall quality, syntactic density, T-unit length, or productivity. With the exception of improvement in editorial usage which generally favored the experimental groups non-significantly, improvements in compositions were randomly distributed between the experimental and control students. Correlational data suggested low to moderate relationships between the two skills, but ranking groups of students on reading and writing scores showed no relationships between the two. This lack of evidence flies in the face of a good deal of previous correlational and descriptive data, a strong *prima-facie* case for a relationship, and over two centuries of theory and speculation.

Failure to prove that there is a causative relationship between reading and writing skills is not the same as proving there is no relationship—it cannot be said that measurable change in reading skill *does not* produce measurable change in writing skill; it can be said only that under the conditions outlined in the current experiment, improvement in reading skill *did not* lead to improvement in writing performance. The complete lack of support for the hypotheses in this study, however, raises serious questions about the entire line of research of which it forms a part. Because virtually no changes in writing performance accompanied the changes in reading skill, it does not appear likely that reading change will lead to any automatic change in writing performance.

Nor can the results of this study be taken to discount the widely held belief that wide reading will improve writing skill. This thesis, too, has a good *prima-facie* case: wide reading brings one into contact with both the more complex structures and the more sophisticated vocabulary of the written dialect. As Veal (1975) points out, writing is, after all, book-talk and one learns book-talk through reading. On the other hand, while studies by Heys (1962) and DeVries (1970) appeared to offer empirical promise for the notion that wide reading improved writing, a carefully controlled five-year study by Elley, Barham, Lamb, and Wyllie (1976) found only small and statistically non-significant advantages at the end of the first year and no advantage thereafter for a group which had two hours of free reading per week in place of two different kinds of grammar instruction offered to other groups.

A certain sterility has marked the research that investigates the potentially dynamic relationship between reading and writing—and the study reported here

shares this disappointing and frustrating character. A few concluding comments about this general line of research may be in order.

The situation is perhaps not quite so bleak as Stotsky (1983) paints it in a recent review of the whole area of reading-writing relationships: "Almost all studies that sought to improve writing through reading instruction were ineffective." Only two studies lay recent claim to having improved reading as stage one of a study directed at the improvement of writing as stage two. These are Matt's study (1977), and the one reported here. Matt did not use standard measures, and neither study succeeded in improving writing. Because the entire line of research remains baffling—in the specific sense that a classic situation prevails where there is substantial correlation but no evidence of a causal relation—there is some risk of undue pessimism unless the issue is kept very precise. Most of the May 1983 issue of *Language Arts*, for instance, is taken up with the question of reading/writing relationships, but there is no shred of hard evidence in it that reading improvement can produce writing improvement. Yet everyone knows, in spite of the lack of evidence, that this almost has to be the case.

What then is wrong with the research? First, very little of it has sought to answer the direct experimental question. Stotsky's (1983) reference to "almost all the studies" is misleading, inasmuch as there have apparently been only two studies in the past 38 years (since Mathews, Larsen, and Butler in 1945) that met the necessary first stage in the experiment—reading improvement—and only eight altogether that even aimed at the question (Belanger, 1978b, p. 73).

Second, none of the three studies (including Mathews) that in fact improved reading as the first step to improving writing dealt adequately with what seems, on reflection, to be the crucial issue of why anyone writes in the first place. The current study assumed that if reading were improved, writing would also automatically improve. But writing is never an "automatic" process. Reading can to some extent intrude itself on a reluctant recipient, as when one sees a sign that says "Danger—road closed." But writing is never automatic in this sense. Therefore, the present study, and the others in the same vein, have been unrealistic in their failure to provide some active incentive for the students to write. In the present study this failure was reflected in the reduced average length of papers on the pre-post comparison. The students knew much better than the researchers that there was, in fact, no persuasive reason for them to write at all on any of the three occasions. Therefore, no real test of writing competence or skill was provided.

The reading treatment may have improved their capacity to write or it may not. But simply telling the students to write the third of three trivial papers without meaningful preparation, audience, or reward is slim evidence indeed of what the students could have done. This sort of purposeless writing has been the hallmark of writing research, and we should all have known better. None of us would dream of writing seriously for the slim rewards we have offered to students in our experiments.

Either reading does not affect writing—which continues to seem most unlikely in the face of the celebrated correlations between the two—or future research will have to abandon the naive assumption that students will go to the daunting effort of writing their best simply to make our experiments come out productively. If students are to write well, the writing will certainly have to be meaningful to them in the first place, and they will have to receive the extrinsic rewards (grades, praise, audience, the possibility of publication, etc.) that are necessary to drive the rest of us to our writing desks. Short of this, serious writing will simply not occur.



Notes

1. To indicate if the subjects chosen for the quality and Syntactic Density Score analyses were representative of the sample population, the initial reading scores of the subjects in these subsamples were compared with those of the subjects in the remainder of the classes. Use of t-tests revealed that differences between the sample and subsamples did not reach the .05 level of confidence, the probabilities falling between .70 and .99.
2. Because of the paucity of evidence supporting the hypotheses of the investigation, only summaries of the findings on the four major and one subsidiary analysis are presented here. An in-depth analysis and discussion of all 36 null hypotheses can be found in the original document (ED 163 409, pp. 265-347).
3. The apparent anomaly of students "losing" quality over the six months can be at least partially attributed to the attitude of some high school students to written work in June: school "officially" continued until June 28 but unofficially appeared to end before June 1. See also concluding comments.
4. The pretest, of course, is unaltered by the treatment. Therefore, one might expect any natural relationships between reading skill and writing skill to be most obvious here.

References

- Ackerman, S. A. (1969). *An analysis of first grade children's writings as they reflect three different methods of teaching initial reading*. Unpublished doctoral dissertation, Chico State College.
- Artley, A. S., Hildreth, G. H., Townsend, A., Beery, A., & Dawson, M. (1954). *Interrelationships among the language arts*. Research Bulletin of the National Conference on Research in English.
- Baker, W. D. (1954). An investigation of characteristics of poor writers. *College Composition and Communication*, 5, 23-27.
- Becker, S. L. (1963). Relationships among the communication skills. *Central States Speech Journal*, 14, 258-263.
- Belanger, J. F. (1978a). Calculating the syntactic density score: A mathematical problem. *Research in the Teaching of English*, 12, 149-153.
- Belanger, J. F. (1978b). *Reading skill as an influence on writing skill*. Unpublished doctoral dissertation, University of Alberta. (ERIC Document Reproduction Service No. ED 163 409)
- Bell, J. B. (1971). *A study of the written composition interests of senior high school students*. Unpublished doctoral dissertation, University of Oregon.
- Braddock, R., Lloyd-Jones, R., & Schoer, L. (1963). *Research in written composition*. Champaign, IL: National Council of Teachers of English.
- Broudy, H. S., & Palmer, J. R. (1965). *Exemplars of teaching method*. Chicago: Rand-McNally.
- Buros, O. K. (Ed.). (1965). *The sixth mental measurements yearbook*. Highland Park, NJ: The Gryphon Press.
- Burrows, A. T., Parke, M. B., Edmund, N. R., Deboer, J. J., Horn, T. D., Herrick, V. E., & Strickland, R. G. (1961). *Children's writing: Research in composition and related skills*. Champaign, IL: National Council of Teachers of English.
- Collins, C. (1981, November). *The use of writing to improve reading*. Paper presented at the annual meeting of the National Council of Teachers of English, Boston, MA. (ERIC Document Reproduction Service No. ED 214 102)
- Combs, W. E. (1977). Sentence-combining practice aids reading comprehension. *Journal of Reading*, 21, 18-24.
- Crowhurst, M. (1980). Syntactic complexity and teachers' quality ratings of narrations and arguments. *Research in the Teaching of English*, 14, 223-231.
- Davis, F. B., & Davis, C. C. (1962). *Davis reading test: Manual, Series 1 and 2*. New York: The Psychological Corporation.



- DeLuca, E. B. (1979). The effect of writing in the expressive mode on the general and specific reading comprehension of underprepared, college-level readers (Doctoral dissertation, Indiana University of Pennsylvania). *Dissertation Abstracts International*, 40, 5343A.
- DeVries, T. (1970). Reading, writing frequency, and expository writing. *Reading Improvement*, 7, 14-15.
- Diederich, P. B. (1957). *The problem of grading essays*. Princeton, NJ: Educational Testing Service (mimeographed).
- Diederich, P. B. (1974). *Measuring growth in English*. Champaign, IL: National Council of Teachers of English.
- Eckhoff, B. (1983). How reading affects children's writing. *Language Arts*, 60, 607-616.
- Elley, W. B., Barham, I. H., Lamb, H., & Wyllie, M. (1976). The role of grammar in a secondary school English curriculum. *Research in the Teaching of English*, 10, 5-21.
- Eurich, A. C. (1931). *The reading abilities of college students: An experimental study*. College Problems Series. Minneapolis, MN: University of Minnesota Press.
- Evanechko, P., Ollila, L., & Armstrong, R. (1974). An investigation of the relationships between children's performance in written language and their reading ability. *Research in the Teaching of English*, 8, 315-326.
- Fishco, D. T. (1966). *A study of the relationship between creativity in writing and comprehension in reading of selected seventh grade students*. Unpublished doctoral dissertation, Lehigh University.
- Fuller, K. M. (1974). An investigation of the relationship between reading achievement and oral and written language of students enrolled in reading and English classes at Gainesville Junior College (Doctoral dissertation, University of Georgia). *Dissertation Abstracts International*, 35, 6692A.
- Fyfe, T. W. (1965). *A study of the effects of the Initial Teaching Alphabet on written composition at Primary III level*, Unpublished thesis, St. Andrews University. (ERIC Document Reproduction Service No. ED 032 295)
- Golub, L. S., & Kidder, C. (1974). Syntactic density and the computer. *Elementary English*, 51, 1128-1131.
- Grobe, C. (1981). Syntactic maturity, mechanics, and vocabulary as predictors of quality ratings. *Research in the Teaching of English*, 15, 75-86.
- Harris, M. W. (1975). *Second grade syntax attainment and reading achievement*. Paper presented at the annual meeting of the International Reading Association, New York, NY. (ERIC Document Reproduction Service No. ED 106 764)
- Harris, T. L. (1969). Reading. In R. L. Ebel (Ed.), *Encyclopedia of educational research* (pp. 1068-1104). Toronto: Collier-Macmillan Canada.
- Harris, W. H. (1977). Teacher response to student writing. *Research in the Teaching of English*, 11, 175-185.
- Heller, M. F. (1980, December). *The reading and writing connection: An analysis of the written language of university freshmen at two reading levels*. Paper presented at the annual meeting of the National Reading Conference, San Diego, CA. (ERIC Document Reproduction Service No. ED 216 342)
- Heys, F., Jr. (1962). The theme-a-week assumption: A report of an experiment. *English Journal*, 51, 320-322.
- Hildreth, G. H. (1954). Interrelationships between written expression and the other language arts. *Elementary English*, 31, 40-48.
- Hill, S. S. (1982). Relationships between reading and writing performance: A correlational study of Metropolitan Reading subscores and National Assessment Writing scores (Doctoral dissertation, University of Florida). *Dissertation Abstracts International*, 43, 1862A.
- Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (Research Report No. 3). Urbana, IL: National Council of Teachers of English.

- Johnson, N. (1976). *A comparison of syntactic writing maturity with reading comprehension*. Paper presented at the annual meeting of the International Reading Association, Texas State Council Conference. (ERIC Document Reproduction Service No. ED 141 794)
- Keltner, J. (1957). The hardest knife. *Speech Teacher*, 6, 275-284.
- Kuntz, M. H. (1975). *The relationship between written syntactic attainment and reading ability in seventh grade*. Unpublished doctoral dissertation, University of Pittsburgh.
- Lazdowski, W. P. (1976). Determining reading grade levels from analysis of written compositions (Doctoral dissertation, New Mexico State University). *Dissertation Abstracts International*, 37, 1504A.
- Levine, S. S. (1976). The effect of transformational sentence-combining exercises on the reading comprehension and written composition of third-grade children (Doctoral dissertation, Hofstra University). *Dissertation Abstracts International*, 37, 6431-32A.
- Loban, W. D. (1963). *The language of elementary school children* (Research Report No. 1).
- Loban, W. D. (1966). *Language ability: Grades seven, eight and nine*. Washington: U.S. Government Printing Office.
- Loban, W. D. (1967). *Language ability: Grades ten, eleven, and twelve*. Washington: U.S. Government Printing Office. Urbana, IL: National Council of Teachers of English.
- MacNeill, T. B. (1982). *The effect of sentence-combining practice on the development of reading comprehension and the written syntactic skills of ninth grade students* (Research prepared at the University of Alberta). (ERIC Document Reproduction Service No. ED 182 465)
- Martin, E. B. H. (1976). *Administration of the S.O.S. reading technique*. Edmonton, Alberta (duplicated).
- Martin, R. G. (1968). *A prediction formula for a sample of "good" writing*. Unpublished doctoral dissertation, University of South Carolina.
- Martin, R. G. (1974). The case of Bruce: A first report on a secondary reading device. *Alberta English*, 14, 17-21.
- Martin, R. G. (1975). S.O.S.: A second report on a secondary reading device. *Alberta English*, 15, 42-44.
- Martin, R. G. (1977). More basic than basics. *Arizona English Bulletin*, 20, 28-33.
- Martin, R. G., & Belanger, J. F. (1975). *A longitudinal assessment of effects of S.O.S. on reading and writing at two grade levels*. Final report to the Edmonton Public School Board, Edmonton, Alberta (duplicated).
- Martin, R. G., & Martin, E. B. H. (forthcoming). *Reading and the subconscious* (tentative title).
- Mathews, E. G., Larsen, R. P., & Butler, G. (1945). Experimental investigation of the relation between reading training and achievement in college composition classes. *Journal of Educational Research*, 38, 499-505.
- Matt, D. W. (1977). An enquiry into empirical relationships between the reading and writing of exposition and argument (Doctoral dissertation, State University of New York at Albany). *Dissertation Abstracts International*, 38, 4631-32A.
- Mazurkiewicz, A. J. (1973). *i.t.a. revisited*. Paper read to the College Reading Association. (ERIC Document Reproduction Service No. ED 083 569)
- McConnell, M. A. (1982). The effect of literature exposure and writing practice on the original narrative writing of second grade children (Doctoral dissertation, Ohio State University). *Dissertation Abstracts International*, 43, 2619A.
- Meckel, H. C. (1963). Research on teaching composition and literature. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 966-990). Chicago: Rand-McNally.
- Miller, B. D., & Ney, J. W. (1968). The effect of systematic oral exercises on the writing of fourth-grade students. *Research in the Teaching of English*, 2, 44-61.



- Mohler, L. E. Z. (1983). The interrelationships of fourth grade students [sic] reading achievement, oral generation, and written production as determined by factors found in four readability formulas (Doctoral dissertation, Montana State University, 1982). *Dissertation Abstracts International*, 43, 2222A.
- Nagle, J. E. (1972). The effects of a directed writing activity in eighth grade social studies instruction on general reading achievement and on social studies reading achievement (Doctoral dissertation, Temple University). *Dissertation Abstracts International*, 33, 1523A.
- Obenchain, A. (1971). Effectiveness of the precis essay question in programming the sequential development of written composition skills and the simultaneous development of critical reading skills (Master's thesis, George Washington University). In S. L. Stotsky, (1975), Sentence-combining as a curricular activity: Its effect on written language development and reading comprehension. *Research in the Teaching of English*, 9, 30-71.
- O'Donnell, R. C. (1976). A critique of some indices of syntactic maturity. *Research in the Teaching of English*, 10, 31-38.
- Palmer, W. S. (1971). Reading, writing and the realm of reason. *Phi Delta Kappan*, 52, 473-475.
- Perry, M. L. (1980). A study of a literary models approach to composition on writing and reading achievement (Doctoral dissertation, Boston University School of Education). *Dissertation Abstracts International*, 40, 6137A.
- Potter, R. R. (1966). *An exploratory study of the relationship between certain aspects of sentence structure and the overall quality of tenth-grade writing*. Unpublished doctoral dissertation, Teachers College, Columbia University.
- Quinn, A. M. V. (1977). Some effects of experience-related reading instruction on the written language development of first-grade children (Doctoral dissertation, University of Maine). *Dissertation Abstracts International*, 38, 1946A.
- Reed, E. (1967). Improving comprehension through study of syntax and paragraph structure in seventh grade classes. In J. A. Figurel (Ed.), *Forging ahead in reading. Proceedings of the International Reading Association*, 12, 575-579.
- Russell, D. H., & Fea, H. R. (1963). Research on teaching reading. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 825-928). Chicago: Rand-McNally.
- Schneider, V. L. (1971). A study of the effectiveness of emphasizing the teaching of reading skills to improve composition skills in remedial English classes at Kansas City Community Junior College. Unpublished doctoral dissertation, University of Kansas.
- Schonell, F. J. (1942). *Backwardness in the basic subjects*. Toronto: Clarke, Irwin.
- Shanahan, T. E. (1980). A canonical correlational analysis of the reading-writing relationship: An exploratory investigation (Doctoral dissertation, University of Delaware). *Dissertation Abstracts International*, 41, 1007A.
- Shapiro, B. J., & Shapiro, P. P. (1973). The effect of reading method on composition: ITA vs TO. *Journal of Reading Behavior*, 5, 82-87.
- Siedow, M. D. (1973). *Relationships between syntactic maturity in oral and written language and reading comprehension of materials of varying syntactic complexity*. Unpublished doctoral dissertation, Indiana University.
- Smith, C. B. (1968). *Relation of first grade reading and composition*. Paper presented to the International Reading Association. (ERIC Reproduction Service No. ED 020 080)
- Stewart, M. F., & Grobe, C. (1979). Syntactic maturity, mechanics of writing, and teachers' quality ratings. *Research in the Teaching of English*, 13, 207-216.
- Stewart, R. W. (1969). Curriculum change as the result of introducing children to reading through ita in Bethlehem. *The ita Journal*, 3, 17-22.
- Stotsky, S. L. (1983). Research on reading/writing relationships: A synthesis and suggested directions. *Language Arts*, 60, 627-642.
- Thomas, F. L. (1976). The extent of the relationship between reading achievement and writing achievement among college freshmen (Doctoral dissertation, University of South Carolina). *Dissertation Abstracts International*, 37, 6320A.



- Thompson, W. N. (1967). *Quantitative research in public address and communication*. New York: Random House.
- Tovatt, A. L., & Miller, E. L. (1967). *Oral-Aural visual stimuli approach to teaching written composition to 9th grade students* (Report No. BR-5-0398, Muncie, Indiana, Burris Laboratory School, Ball State University). (ERIC Document Reproduction Service No. ED 015 204)
- Veal, R. (1975, December 8). Cited in M. Sheils, Why Johnny can't write. *Newsweek*, 86, 59.
- Weaver, P. A. (1977). *Component skills in reading comprehension: An investigation of the effects of sentence organization instruction*. Paper presented at the annual meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. ED 136 241)
- Yusuf, D. R. (1982). Reading performance test scores and the impact of a writing program (Doctoral dissertation, Temple University). *Dissertation Abstracts International*, 43, 745A.
- Zanotti, R. J. (1970). A study of the use of the tape recorder as an aid to written composition at the sixth-grade level (Doctoral dissertation, State University of New York at Buffalo). *Dissertation Abstracts International*, 31, 1520.

WINSTON E. RAMPAUL

*University of Manitoba*

MOTIE SINGH

*Frontier School Division*

and

JOHN DIDYK

*Manitoba Department of Education*

## The Relationship Between Academic Achievement, Self-Concept, Creativity, and Teacher Expectations Among Native Children in a Northern Manitoba School

*There is a dearth of empirical evidence pertaining to Native learners in northern Manitoba communities, especially with respect to significant variables, namely academic achievement, self-concept, creativity, and teacher expectations. In this study, the researchers focused on determining whether significant positive relationships existed among these variables with a grade three class of twenty-two children, a grade four class of nineteen children, and with the total group of forty-one children in grades three and four combined. To measure the variables the following instruments were used: for academic achievement, the CTBS; for self-concept, the Michigan State General Self-Concept of Ability; for teacher expectation, the Teacher Rating of Pupil Attitude Scale; and for creativity, the Torrance Tests of Creative Thinking. Significant positive correlations beyond the .05 level, ranging from .35 to .61, were found for grades three and four among self-concept, academic achievement, and teacher expectations. Only one significant positive correlation was discovered (.38) for creativity and teacher expectations. Support for the phenomena of age-grade deceleration and academic retardation with increasing age was discovered. Recommendations for ameliorating the unsatisfactory condition were made and further cross-validation study was suggested.*

---

Dr. Winston Rampaul holds a PhD from the University of Wisconsin, Madison and is now an Associate Professor in Educational Psychology at the Faculty of Education, University of Manitoba.

Mr. Motie Singh has a Master's degree in Educational Psychology from the University of Manitoba and is now a resource teacher in Manitoba.

Dr. John Didyk obtained his PhD at the University of Manitoba and is presently a program supervisor with the Manitoba Provincial Department of Education.

There is much research evidence to suggest that academic achievement correlates positively with self-concept (Brookover, Schweiter, Beady, Flood, & Wisembaker, 1976; Caplin, 1969; Purkey, 1970; Stipek & Hoffman, 1980), creativity (Getzels & Jackson, 1962; Guilford, 1968, 1977; Schempp, 1981) and teacher expectations (Braun, 1977; Brophy & Good, 1970; Cooper, 1979; Good & Brophy, 1977; Rosenthal & Jacobson, 1968). However, previous research has not investigated these relationships with Native Indian children. This study sought to examine the interrelationships between academic achievement, self-concept, creativity, and teacher expectations among Native students in a small elementary school, in a Northern Manitoba community.

Empirical evidence pertaining to the relationships among these important variables will serve to clarify the nature of these concepts, namely academic achievement, teacher expectations, self-concept, and creativity. In turn, knowledge generated about these will extend educators' understanding and insight which will eventually result in the improvement of instructional sequencing and management. Hopefully, academic deceleration and retardation of Native children will decline.

### *Background*

There is a dearth of research pertaining to academic achievement involving Native children, in particular as related to children's self-concept, creativity, and teacher expectations. However, research findings from studies involving non-Native subjects suggest high positive inter-correlations among these variables, and these may have implications for the education of Native children.

### *Self-Concept*

Self-concept, broadly defined, is a person's perceptions of himself/herself. These perceptions are formed through one's experience with and interpretations of one's environment, and are influenced especially by reinforcements, evaluations by significant others, and one's attributions for one's own behavior (Shavelson, Hubner, & Stanton, 1976).

Considerable research evidence exists to suggest a relationship between self-concept and academic achievement. Hansford and Hattie (1982) used meta-analysis to examine the relationship between various self measures and measures of performance and achievement; the statistical results of 128 studies were transformed into a common measure, correlation coefficients. The studies represented a total sample of 202,823 persons and produced a data base of 1,136 correlations between self-ratings and performance measures. A range in the relationship of  $-.77$  to  $.96$  was reported with an "average" correlation of  $.21$ . It was found that this average relationship was modified by a number of variables which included: grade-level of subjects, socioeconomic status, ethnicity, ability of subjects, and type of performance/achievement measures.

Self-concept changes may be regarded as a change of attitude towards the subject and not a change in academic ability. At the elementary level, if a positive attitude towards learning can be fostered in students, it seems likely that a positive self-concept of ability would result. Research conducted with Indian children in Northern communities (Kirkness, 1978; Lenton, 1979) has shown that there is an age-grade retardation that begins at grade one and tends to increase as Indian students progress through school. Since academic achievement has been found to correlate positively with self-concept (Brookover, et al., 1976; Caplin, 1969; Stipek &



Hoffman, 1980) it might be assumed that as the Indian child progresses through the school, his/her self-concept and academic achievement would tend to decrease simultaneously.

Many researchers (Alvord & Glass, 1974; Bacher, 1965; Deo & Sharma, 1970) have investigated the relationship between academic achievement and self-concept at the secondary level. At the secondary level, however, the individual has probably already associated his/her self-concept of ability with his/her success or failure in that subject area. Hence, this research focused on the early elementary years for Native Indian children. As suggested by Beaird (1965), if low self-concept can be identified and found to be related to academic achievement during the early elementary years, then attempts can be made to enhance the individual's self-concept of ability. This may alter his/her attitude to education before he/she becomes a potential failure or dropout.

### *Creativity*

There appears to be a low positive relationship between creativity and academic achievement (Getzels & Jackson, 1962; Guilford, 1968; Torrance, 1965). Torrance's (1965) research suggested that low correlation occurred because the top twenty percent of subjects who scored high on measures of intelligence were not the top twenty percent of the subjects who scored high on measures of creativity. According to Guilford (1968), traditional intelligence tests tended to require convergent thinking while creativity required divergent thinking.

The uncertainty and lack of consensus about the essence of creativity and the lack of success in deriving a satisfactory operational definition constitute a fundamental problem, particularly with respect to measurement. Definitional disagreement persists. For example, creativity has been defined as an ability involving fluency, flexibility, and originality (Guilford, 1959), as the forming of associative elements into new combinations (Mednick, 1962), as a novel product (Stein, 1968), as moments of sudden, surprising insight (Bruner, 1962; Piaget, 1971), and as the process of bringing something new into existence (Johnson & Johnson, 1975). There is also disagreement on whether it is an inherited trait or a process anyone can learn (Johnson, 1979). Studies, such as this one, continue the search for greater understanding and relatedness which will help in firming up a more widely acceptable definition.

Results of studies suggest that the most creative children are not necessarily the most intelligent nor are they those who score best on achievement tests. Children of lower-class income groups have been found to be more creative than children from middle-class income groups (Gezi, 1969). Since most Native Indian children come from a low-income group, they may be generally creative; further, Torrance's (1965) and Guilford's (1968) findings suggest that the more creative Native children may not necessarily be the top achievers.

### *Teacher Expectations*

According to Schmuck and Schmuck (1983), an interpersonal expectation is a prediction of how another person will behave in a particular social setting, and may develop through the individual's own personality structure, his/her generalized images of what the other person is like, or as a result of specific reactions perceived to be coming from the other (p. 68). The effects of teachers' expectations on students have been found to reflect what has been described as the "self-fulfilling prophecy,"

which occurs when a “false” definition of a situation evokes a new behavior which makes the originally “false” conception come true (Schmuck & Schmuck, 1983, p. 90).

The classic study by Rosenthal and Jacobson (1968) demonstrated that a teacher’s expectation of his/her pupils’ intellectual competence can serve as an educational self-fulfilling prophecy. Although Rosenthal and Jacobson’s research has been severely criticized for its methodology and statistical analysis, follow-up research (Brophy & Good, 1970; Cooper, 1979; Good & Brophy, 1977; Rist, 1970) has found evidence to support the contention that teachers’ expectations affect teacher behavior toward the student, student behavior, and student achievement. According to Schmuck and Schmuck (1983), such reviews/researches leave little doubt that the expectations of teachers have important and real effects on students. Teachers’ expectations appear to influence achievement, self-esteem, and the satisfaction of students (p. 94).

Rist (1970) found that teacher expectation was influenced by the socio-economic status of the students; subjects from lower socio-economic backgrounds tended to be placed into lower curriculum tracks. Since most Native Indian children tend to come from a lower-class income group, teacher expectation may be an important variable in influencing both the self-concept and the academic achievement of Indian children. Studies by Brookover, Erickson, and Joiner (1967) and Clifton (1972) have found that teacher expectation influences students’ self-concept. The relationship between teacher expectation, self-concept, and academic achievement appears to be an intimate one.

### *The Problem*

Research is needed to determine the relationships among the variables of academic achievement, self-concept, teacher expectation, and creativity of elementary-level Native Indian children in a school in a Northern Manitoba community. The term “Native” refers to Canadians of aboriginal descent and can include status and non-status Indians, Inuit, and Metis.

The basic question is, to what extent do these variables relate to each other? On the evidence of previous research, high positive inter-correlations should exist between teacher expectation, self-concept of ability, and academic achievement. The relationships among these variables and creativity should be less striking.

### *Significance of the Problem*

Age-grade retardation and progressive academic deceleration associated with increasing age for Native children have been recorded in the research literature. More information about the nature of academic achievement and its related variables is needed in order to provide greater understanding, insights, and clues about ways in which academic achievement of Native children could be facilitated. Such information could be useful in removing misconceptions, furnishing hope and optimism, and forming a scientific and objective basis for meaningful intervention. This study is aimed at generating much-needed information about academic achievement with respect to Native children. The results of the study can serve as a useful starting point for a much larger study. Information generated could prove to be helpful to both pre-service and in-service teachers of Native children.

More specifically, professors of teacher education should ensure that all in-service and pre-service teachers are aware of the basic research findings concerning the



development of a negative self-concept. Attention should also be given to the trap of low teacher expectations of their students, and the stifling of creativity by teacher behavior.

### *Method*

#### *Population and Sample*

The study was carried out in a school situated in Northern Manitoba, about 900 miles from Winnipeg. There were 189 students enrolled at the school with a staff of 14 teachers, including two teacher aides. Ninety-seven percent of the school population were Cree speaking, hence English was a second language for most students. The remoteness of the community, lack of recreational facilities, and lack of professional personnel, might allow the community to be seen as a disadvantaged environment for both the Euro-Canadians and the Indians.

The subjects of the study were 22 students of grade three (12 boys and 10 girls) and 19 students of grade four (7 boys and 12 girls), together with the respective classroom teachers. These 41 subjects were chosen because the teachers were willing to permit them to participate.

#### *The Instruments*

*A Measure of Academic Achievement.* Academic achievement was measured by the Canadian Test of Basic Skills (CTBS), Form 3, Levels 9-14, (1976). These tests are designed to assess levels of skills development in reading, vocabulary, the mechanics of written language, methods of study, and mathematics for students in grades three to eight. Reliability for the CTBS varies from test to test and from grade to grade. Internal consistency reliability coefficients range from .87 to .96, while composite reliability ranges from .97 to .98 for all grades (CTBS Manual, 1982). Kuder-Richardson (K-R21) reliability coefficients were calculated for use in this research; coefficients of .89 and .41 were obtained for grades three and four, respectively. The low reliability at grade four was probably due to the small number of subjects and the narrow range of marks.

Eleven subtests generated five scores; three for Language Arts, and one each for Mathematics and Work Study Skills. Academic Achievement was represented by the composite scores obtained from these five scores.

*A Measure of Creativity.* Creativity was measured by the Torrance Test of Creative Thinking (TTCT), Thinking Creatively with Pictures (Figural Tests, Forms A and B) (1966). The TTCT was designed to measure four aspects of "creative thinking"—fluency, flexibility, originality, and elaboration. The two forms of this test, A and B, include three activities: 1) picture construction activity, yielding scores for elaboration and originality; 2) incomplete figures activity, and 3) repeated figures activity, each providing scores for fluency, flexibility, originality, and elaboration. Creativity was measured by the accumulated score on the above four abilities. Test-retest reliabilities for the TTCT range from .50 to .93 over one to two-week periods, and from .35 to .73 over three-year periods (Buros, 1972). Reliability coefficients obtained for this study were: 1) Spearman-Brown reliability coefficients, .82 and .94 for grades three and four respectively; and 2) K-R21 reliability coefficients, .95 and .97, respectively.

*A Measure of Self-Concept of Ability.* The Michigan State General Self-Concept of Ability Scale (SCA) was used as a measure of self-concept. Other scores used were: the Perceived Parental Evaluation of Ability (PPEV), the Perceived



Friends' Evaluation of Ability (PFEV), and Perceived Teachers' Evaluation of Ability (PTEV). These scales determine the effects of a student's perceptions of his academic ability as seen by significant others—parents, friends, and teachers. A high numerical score on these scales represents a more positive self-concept, while a lower score represents a less positive self-concept.

The Hoyt's analysis of variance reliability coefficients of the SCA range from .77 to .88 for grade seven through grade twelve, while reliability for the PPEV, PFEV, and PTEV range from .75 to .92 (Brookover et al., 1967, p. 61). Spearman-Brown reliability coefficients calculated for this research were .87 and .84 for grades three and four respectively, while K-R21 reliability coefficients were .96 and .91 respectively.

*A Measure of Teacher Expectation.* The Teacher Rating of Pupil Attitude Scale (TRPA) developed at the Tennessee Technology University (1969), assessed teacher expectation of students' attitudes and ability in academic achievement and creativity. The TRPA required teachers to respond to sixteen questions. The TRPA was considered an appropriate measuring instrument for this research because it was originally designed for use with culturally disadvantaged children and minority groups who had enrichment programs to develop their creative expressions and self-concept, and this study involved the academic achievement, self-concept, and creativity of subjects from a similar background. Reliability coefficients for the TRPA are not available, but K-R21 reliability coefficients estimated for this study were .83 and .93 for grades three and four, respectively.

#### *Administration and Scoring*

All the tests were administered in accordance with the developers' specified directions. The CTBS was administered over a four-day period in both grades three and four classrooms. Each of the four testing sessions lasted from 65 to 80 minutes. The TTCT was administered in the week following the testing on the CTBS. The TTCT Figure Tests, Form A and Form B, were administered on separate days. The SCA, PPEV, PFEV and PTEV were administered on the same day; while the TRPA was done by the classroom teachers during the final week of the school year. For each student four scores were computed: one each for academic achievement, self-concept, creativity, and teacher expectation.

#### *Results and Discussion*

Twenty-two sets of scores were computed for grade three students and nineteen sets for grade four students. Data were analyzed descriptively (means, standard deviations, correlation coefficients) for each grade according to sex, and all the children as a single group. Pearson product moment correlations were computed to examine the relationships between pairs of variables. Scores were analyzed using the ST32 computer program run on the AMDAHL 470/V7 computer at the University of Manitoba. Tests of statistical significance were conducted at the .05 level.

Table 1 summarizes the means and standard deviations of scores on each variable: self-concept, creativity, teacher expectation, and academic achievement. Except for sex differences in creativity, none of the differences between pairs of means for the four variables were found to be statistically significant. Boys appeared to perform more creatively than girls (303.74 vs. 297.26).

It is interesting to note that grades three and four, generally speaking, had approximately the same means and standard deviations on all the variables with

slight exceptions. For example, grade four scored only about five points more on self-concept than grade three. Table 1 shows the grade three mean for self-concept to be 88.68 and grade four 93.11. Only the grade three creativity mean was higher than the grade four creativity mean. On all other variables the means for grade four were higher and were substantially so for academic achievement: 131.42 for grade four and 126.59 for grade three.

TABLE 1  
MEANS AND STANDARD DEVIATION OF SCORES ON SELF-CONCEPT,  
CREATIVITY, TEACHER EXPECTATION, AND ACADEMIC ACHIEVEMENT  
FOR ALL THE SUBJECTS AS A GROUP, FOR EACH GRADE LEVEL  
AND WHEN GROUPED BY SEX

| TESTS                | GRADE  |        |        |                |                 |
|----------------------|--------|--------|--------|----------------|-----------------|
|                      | 3 & 4  | 3      | 4      | 3 & 4          | 3 & 4           |
|                      | (N=41) | (N=22) | (N=19) | Boys<br>(N=22) | Girls<br>(N=19) |
| Self-Concept         |        |        |        |                |                 |
| Mean                 | 90.73  | 88.68  | 93.11  | 89.68          | 91.96           |
| S.D.                 | 14.55  | 15.99  | 12.69  | 17.70          | 10.10           |
| Creativity           |        |        |        |                |                 |
| Mean                 | 303.15 | 304.45 | 301.63 | 303.73         | 297.26          |
| S.D.                 | 88.59  | 78.49  | 103.17 | 75.04          | 111.23          |
| Teacher Expectation  |        |        |        |                |                 |
| Mean                 | 59.71  | 58.86  | 60.68  | 60.55          | 58.74           |
| S.D.                 | 10.86  | 9.12   | 12.78  | 10.69          | 11.27           |
| Academic Achievement |        |        |        |                |                 |
| Mean                 | 128.83 | 126.59 | 131.42 | 131.41         | 125.84          |
| S.D.                 | 22.27  | 28.31  | 12.38  | 18.08          | 26.51           |

However, these mean academic achievement scores of the grade three (126.59) and four (131.42) subjects were below their grade equivalents when transformed to the norms of the CTBS: 2.5 vs. 3.9 or greater, a difference of 1.4 for grade three subjects; and 2.6 vs. 4.9 or greater, a difference of 2.3 for grade four subjects. Since the academic achievement test was administered towards the end of the school year, these results tend to support previous research findings attesting to an age-grade deceleration in achievement for Native children in Canada (Kirkness, 1978; Lenton, 1979).

Table 2 provides a summary of the inter-correlations found among the four variables. Significant positive correlations were found between: 1) self-concept and teacher expectation for the total sample and for boys; 2) self-concept and academic achievement across all grade levels and for both sexes; 3) creativity and teacher ex-

pectation and academic achievement across all groups. In general, there was a low correlation between creativity and academic achievement for boys, and the relationship was negative. Correlations between self-concept and creativity were generally low; in most cases, the relationship was negative.

TABLE 2  
SUMMARY OF INTER-CORRELATIONS AMONG THE FOUR VARIABLES

| VARIABLES   | GRADE           |             |             |                         |                                       |
|---|-----------------|-------------|-------------|-------------------------|---------------------------------------|
|   | 3 & 4<br>(N=41) | 3<br>(N=22) | 4<br>(N=19) | 3 & 4<br>Boys<br>(N=22) | 3 & 4<br>Girls<br>(N=19) <sup>a</sup> |
| Self-Concept<br>and Creativity                        | -.01            | -.07        | .06         | -.11                    | .09                                   |
| Self-Concept<br>and Teacher<br>Expectation            | <u>.35</u>      | .36         | .35         | <u>.46</u>              | .21                                   |
| Self-Concept<br>and Academic<br>Achievement           | <u>.48</u>      | <u>.51</u>  | <u>.43</u>  | <u>.61</u>              | <u>.48</u>                            |
| Creativity<br>and Teacher<br>Expectation              | .19             | <u>.38</u>  | .08         | -.04                    | .29                                   |
| Creativity<br>and Academic<br>Achievement             | .08             | .11         | .06         | -.14                    | .19                                   |
| Teacher<br>Expectation<br>and Academic<br>Achievement | <u>.50</u>      | <u>.64</u>  | <u>.45</u>  | <u>.52</u>              | <u>.49</u>                            |

<sup>a</sup> Number in parentheses indicate the number of subjects who completed all the tests.

Note: Coefficients underlined indicate correlations significant beyond .05 level.

Certain limitations related to the instruments and the definitions of variables used should be noted when interpreting the results of this research. Measurement of self-concept was limited to twenty-three statements which could be understood by subjects. These statements may comprise only a small segment of an individual's total self-concept of ability. Further, only aspects of creativity which could be expressed by pictures were measured.

Academic achievement was limited to achievement in three subject areas—Language Arts, Mathematics, and Work Study Skills, which may not be sufficient to make a generalized statement of the subjects' total achievement. However, since these subject areas form the basis for the other subject areas, data on these subjects may still provide a fair assessment of academic achievement.

Finally, only sixteen statements concerning the teacher's expectations of a student's attitude and performance were employed in this research. It seems doubtful that these statements covered every aspect of the teacher's expectations.



On the basis of the findings of this research, one may conclude that Native children at the grades three and four level who possess a high self-concept of ability, and for whom teachers hold high expectations, also tend to be high achievers. These findings support previous research attesting to positive relationships between academic achievement and self-concept (Brookover, et al., 1976; Caplin, 1969; Hansford & Hattie, 1982), and academic achievement and teacher expectation (Braun, 1977; Brophy & Good, 1970; Rosenthal & Jacobson, 1968). The self-concept of Native children also appears to be highly correlated with teacher expectation, supporting conclusions by Brookover et al. (1967) and Clifton (1972). The high achievers among Native students at the grades three and four level were not necessarily the most creative, as evidenced by the low correlation between academic achievement and creativity; in some cases, the relationship appears to be negative.

### *Recommendations*

Discovering repeatedly that age-grade deceleration and academic retardation exist in Native learners, or for that matter, among any learners, should not be construed as an end in itself. Possible ways of resolving the problem should also be identified so that steps might be taken to ameliorate this vexing condition, which results in social injustices and inequality in opportunities for learners. Some of the recommendations that follow might be helpful.

1. Professors of education and teacher trainers should bring to the attention of in-service and pre-service teachers of Native learners the need to be open toward the learning potential of these children rather than stereotyping them as low achievers and somehow communicating low expectations to them either explicitly or through subtle gestures. The available research should be made known and discussed in spite of its limitations and weaknesses.

2. The school and teachers of any community cannot and should not accept the total responsibility for the all-round educational development of Native learners. The education of culturally different or disadvantaged learners should be a team effort, including the family, the community, the church, the media and all those models who come into contact with the learner in his/her formative years.

Educating the culturally different or disadvantaged has become an increasingly paramount issue over the last fifteen years. The Native learners in this study live in, and are educated in, an isolated northern community. They do not come into contact with many of the economic and social privileges that are recognized to be the "carriers" of the salient ingredients that could unlock their potential ability. Home, environmental and school factors have often been cited as influential determinants of school achievement. For example, Score and Weinburg (1976) have mentioned a good home environment, opportunities for language and number skills at home, and a concentrated environmental enrichment as being imperatives for the offsetting of cultural disadvantage. Stone and Nielsen (1982) have identified such crucial extra-school variables as library facilities, books and periodicals, learning supplies (including auto-instructional devices), social support, opportunities for enlarging vocabulary, and exposure to models of good language usage. For school factors, deliberate efforts have to be made to accommodate children during the transitional periods of acculturation, and a careful study of the dynamics and inter-play of the variables involved in school learning is essential.

3. Admittedly, the phenomena of self-concept, teacher expectations, and creativity seem to be concomitants of academic achievement and intellectual growth. Per-

haps a strong desire to resolve the problem of academic retardation could result in educators attacking the condition directly. Such remediation would include considerations for learner entry behaviors, identification of precise behavioral objectives, prescriptive teaching, continual monitoring, and measurement. Undoubtedly, these competencies demand specialized training and education of teachers of Native learners.

There is no shortage of exemplary individualized programs that could prove to be very helpful. Houston (1974) has suggested a competency-based program encompassing precise objectives, performance criteria instructions pertinent to the criteria, and learner accountability in terms of the criteria. Davies (1973) has developed performance-based activity packages as part of a teacher's design to achieve pre-determined learning outcomes through appropriately planned individualization of instruction. Gilstrap and Martin (1975) have recommended drill and practice to ensure specific instructional outcomes and obtain increasing student interest and enjoyment. An alternative approach, suggested by Blackie (1967), seems quite plausible for Native learners. Among other things, Blackie prescribed provisions for individualized approaches through a variety of methods and organizational patterns. He placed emphasis on learners making observable gains and improving their self-management through the manipulation of concrete materials and interaction with each other as well as with the teacher.

Other reinforcement-based programs such as precise teaching (Johnson, 1971), precision teaching (Sprinthall & Sprinthall, 1981), mastery learning (Bloom, 1974), criterion-referenced teaching (Popham, 1973), diagnostic and prescriptive teaching and programming (Stellern, Vasa, & Little, 1976; Woolfolk & McCune-Nicolich, 1984), and DISTAR (Engelman, et al., 1975) have incorporated excellent features of high quality instruction, customized especially for children with special needs. The principles underlying these programs are essentially successful and corrective experiences (Glasser, 1969). Others such as Glaser (1962), Gronlund (1973), Hartman (1975), and Airisian (1971) have specified prescriptions for designing instruction systematically to ensure success and corrective experiences.

4. For greater validity and generalization, further research seems necessary. A more elaborate study should be designed that would focus on a larger group of Native learners, probably in more than one community, and a white Euro-Canadian group of learners. In the meantime, serious consideration should be given to the results of this study in order to avert continued growth of the problem.

Failure to take action may result in the continued increase of social injustices, inequalities of opportunities, potential for racial conflict and disharmonies, and finally a denigrating economic dependence. William Glasser (1969) has stated in his book *Schools Without Failure* that lack of success in schools result in children becoming lonely, delinquent, worthless, and coming to feel that they are not loved. They are separated from the world of the successful by a thick wall. Perhaps, this thick wall for the Native learners is a school system with teachers who don't expect them to do well, who thus help them to experience a low self-concept, who may squelch their creativity, and finally produce academic failure.

#### References

- Airisian, P. W. (1971). The role of mastery learning. In J. H. Block, *Mastery learning* (p. 485). New York: Holt, Rinehart & Winston.



- Alvord, D., & Glass, L. W. (1974). Relationship between academic achievement and self-concept. *Science Education*, 58(2), 175-179.
- Bacher, J. H. (1965). The effects of special class placement on self-concept, social adjustment, and reading growth of slow learners (Doctoral dissertation, New York University, 1964). *Dissertation Abstracts International*, 25(12), 7071.
- Beaird, R. E. S. (1965). Self-concept as related to adolescent school dropouts (Doctoral dissertation, University of Nebraska Teachers' College, 1964). *Dissertation Abstracts International*, 25(10), 5724.
- Blackie, J. (1967). *Inside the primary school*. London: Her Majesty's Stationery Office.
- Bloom, B. S. (1974). Time and learning. *American Psychologist*, 29, 682-688.
- Braun, C. (1977). Pygmalion in the reading circle. *Academic Therapy*, 12(4), 445-543.
- Brookover, W. B., Erickson, E. L., & Joiner, L. M. (1967). *Self-concept of ability and school achievement, III*. (Cooperative Research Project No. 2831). East Lansing, MI: Michigan State University.
- Brookover, W. B., Schweitzer, J. H., Beady, C. H., Flood, P. K., & Wisembaker, J. M. (1976). *Elementary school social climate and school achievement*. (ERIC Document Reproduction Service No. ED 131 602)
- Brophy, J., & Good, T. (1970). Teachers' communications of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology*, 61, 365-374.
- Bruner, J. S. (1962). The conditions of creativity. In H. Gruber, G. Terrell, & M. Wartheimer (Eds.), *Contemporary approaches to creative thinking*. Englewood Cliffs, NJ: Prentice-Hall.
- Buroš, O. K. (Ed.). (1972). *The seventh mental measurements yearbook (Vol. I)*. New Jersey: The Gryphon Press.
- Caplin, M. D. (1969). The relationship between self-concept and academic achievement. *The Journal of Experimental Education*, 37(2), 13-16.
- Clifton, R. A. (1972). *Self-concept and attitudes: A comparison of Canadian Indian and non-Indian students*. Paper presented at the 14th annual meetings of the Western Association of Sociology and Anthropology, Calgary.
- Cooper, H. M. (1979). Pygmalion grows up: A model for teacher expectation communication and performance influence. *Review of Educational Research*, 49, 389-410.
- Davies, I. K. (1973). *Competency-based learning: Technology, management, and design*. New York: McGraw-Hill.
- Deo, P., & Sharma, S. (1970). Self ideal discrepancy and school achievement. *Adolescence*, 5(19), 353-360.
- Engelmann, S., Osborn, J., & Engelmann, T. (1978). *DISTAR*. Chicago: Science Research Associates.
- Getzels, J. W., & Jackson, P. W. (1962). *Creativity and intelligence: Explorations with gifted students*. New York: John Wiley & Sons.
- Gezi, K. I. (1969). *Analyses of certain measures of creativity and self concept and their relationships to social class*. California Educational Research Association. (ERIC Document Reproduction Service No. ED 031 533)
- Gilstrap, R. L., & Martin, W. R. (1975). *Current strategies for teachers: A resource for personalizing instruction*. Pacific Palisades, CA: Goodyear.
- Glaser, R. (1962). Psychology and instructional technology. In R. Glaser (Ed.), *Learning research and education*. Pittsburgh: University of Pittsburgh Press.
- Glasser, W. (1969). *Schools without failure*. New York: Harper & Row.
- Good, T. L., & Brophy, J. E. (1977). *Educational psychology: A realistic approach*. New York: Holt, Rinehart & Winston.



- Gronlund, N. E. (1973). *Preparing criterion-referenced tests for classroom instruction*. New York: Macmillan.
- Guilford, J. P. (1959). Three faces of intellect. *American Psychologist*, 14, 469-479.
- Guilford, J. P. (1968). *Intelligence, creativity, and their educational implications*. California: Robert R. Knapp.
- Guilford, J. P. (1977). *Way beyond the I.Q.* New York: Creative Education Foundation.
- Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research*, 52, 123-142.
- Hartman, C. L. (1975, January). Describing behavior: Search for an alternative to grading. *Educational Leadership*, 274-277, 467.
- Houston, W. R. (1974). *Exploring competency-based education*. Berkeley, CA: McCutchan.
- Johnson, D. W. (1971). The effectiveness of role reversal: The actor or the listener. *Psychological Reports*, 28, 275-282.
- Johnson, D. W. (1979). *Educational psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Johnson, D. W., & Johnson, F. (1975). *Joining together: Group theory and group skills*. Englewood Cliffs, NJ: Prentice-Hall.
- Johnson, E. C. (1973). Precision teaching helps children learn. *Teaching Exceptional Children*. Unpublished manuscript, Western Illinois University, 328.
- Kirkness, V. J. (1978). *Education of Indians in federal and provincial schools in Manitoba*. Evaluation report. Ottawa: Department of Indian Affairs and Northern Development.
- Lenton, S. D. (1979). *The education of Indian children: Long Plain, Dakota Plains, Dakota Tipi Indians, Manitoba, 1965-1979*. Manitoba.
- Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, 69, 220-232.
- Piaget, J. (1971). The theory of stages in cognitive development. In D. Green, M. Ford, & G. Flower (Eds.), *Measurement and Piaget* (pp. 1-11). New York: McGraw-Hill.
- Popham, W. J. (1973). *Criterion-referenced instruction*. Belmont, CA: Fearon.
- Purkey, W. W. (1970). *Self-concept and school achievement*. Englewood Cliffs, NJ: Prentice-Hall.
- Rist, R. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard Educational Review*, 40, 411-451.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart & Winston.
- Schempp, P. E. (1981). Influence of decision-making on attitudes, creativity, motor skills, and self-concept in elementary children. *Dissertation Abstracts International*, 42(6), 2563A.
- Schmuck, R. A., & Schmuck, P. A. (1983). *Group processes in the classroom* (4th ed.). Dubuque, IA: Wm. C. Brown.
- Score, S., & Weinburg, R. A. (1976). I.Q. test performance of black children adopted by white families. *American Psychologist*, 31, 726-739.
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46, 407-442.
- Sprinthall, R. C., & Sprinthall, N. A. (1981). *Educational psychology: A developmental approach* (3rd ed.). Reading, MA: Addison-Wesley.
- Stein, M. (1968). *The creative individual*. In manuscript.
- Stellern, J., Vasa, S. F., & Little, J. (1976). *Introduction to diagnostic-prescriptive teaching and programming*. Glen Ridge, NJ: Exceptional Press.
- Stipek, D. J., & Hoffman, J. M. (1980). Children's achievement-related expectancies as a function of academic performance, histories and sex. *Journal of Educational Psychology*, 72(6), 861-865.

- Stone, D. R., & Neilsen, E. C. (1982). *Educational psychology: The development of teaching skills*. New York: Harper & Row.
- Torrance, E. P. (1965). *Rewarding creative behavior*. New Jersey: Prentice-Hall.
- Torrance, E. P. (1966). *Torrance tests of creative thinking*. New Jersey: Personnel Press.
- Woolfolk, A. E., & McCune-Nicolich, L. (1984). *Educational psychology for teachers*. Englewood Cliffs, NJ: Prentice-Hall.

MARY T. NIXON

and

CHESTER S. BUMBARGER

*University of Alberta*

## Prospective Teachers: Views of Programs and Needs

*Practising Alberta educators were asked to assist in identifying societal and educational trends and to assess their impact on the future role of the teacher. In a second study prospective teachers at the University of Alberta were asked: (1) to give their opinions with respect to the identified societal and educational trends, and (2) to respond in terms of their professional preparation programs as to how well they were being prepared for the identified changes in the role of the teacher in the future.*

*Less than half of the prospective teachers (42 percent) felt adequately prepared to cope with identified societal changes but 54 percent thought that they would be able to adjust to the changing educational context. While 71 percent of the student group perceived their own abilities as "above average" in comparison to their peers, only 30 percent of students would rate their preparation programs as better than "adequate."*

*Students would place a heavy emphasis on the practicum component of their preparation programs. However, reliance on student opinion is a necessary input but not sufficient guide to improvement in teacher education programs; other considerations as well must be weighed in developing program modifications.*

Public schools are clearly servants of their societies. By extension then, it could be argued that teacher preparation institutions are also directed by that same society's views. Teacher preparation programs are largely designed to fit teachers into the existing public education reality. As the society changes, a concurrent necessity is adjustment in, or adaptation of, teacher preparation programs.

---

Mary T. Nixon holds a PhD in Educational Administration and is Associate Research Professor, Educational Administration, at the University of Alberta.

Chester S. Bumbarger holds a DEd in Educational Administration and is Professor of Educational Administration at the University of Alberta.



Changes in the broad societal context—such as the changes in family structure, the technological imperatives of the information society and the economic downturn—are affecting all educational institutions (Toffler, 1981; Bernstein, 1983; Gold, 1983). The public school does not remain unaffected by changes in this larger environment for these forces impinge upon the immediate educational context and demand adjustment. At a time when public confidence in educational institutions seems to be decreasing (Bakalis, 1983) it becomes increasingly necessary for educational institutions to be responsive to pressures for change and to the many publics they purport to serve.

As social change becomes more obvious, reform demands directed at the school begin to surface (Gideonse, 1982; Goodlad, 1982). As these demands become the popular voice, pressures are generated to change the operations of the educational system. These demands are the articulation of a perceived need.

Changes in the educational context may be essentially complete within themselves or may simply foreshadow events not yet on the horizon of education. The degree to which potential changes are seen as “future” is an important judgement since it determines the perceived need for action as contrasted with reflection.

A defined future requires action but attempting to extrapolate from the present in order to predict and plan for a long-term future is probably unproductive. Experience has already underlined our limited ability to calculate the unforeseen variables and the unanticipated events which influence the future (Ravitch, 1983; Apple, 1983). It is nonetheless true that the antecedents of the future do lie in today's decisions. The forces which will influence future conditions are being generated today.

There is certainly a case to be made for viewing the future as “now” with respect to the apparent immediacy of threats to education. Indeed it may not truly be possible to separate a set of “now” activities from a set of “future” activities. Anyone involved with teacher preparation will recognize the inevitable time lag between the start of a new program and the graduating class.

In a society which has become accustomed to an accelerated pace of social change, the intervening time period which separates a present set of conditions from a future set of conditions has become relatively brief (Moore, 1963); indeed any such separation of time may be simply illusory (Toffler, 1981). This represents a severe shortening of available time to identify problems and pose alternative solutions. Such a viewpoint leads naturally to concerns about decisions and changes now that are important in their own rights and not simply as temporary expedients.

Concentration upon “now” problems is at least an acceptable, and may possibly be the only, effective tactic available to educators. With this mind-set, an examination of the current context for education and the potential threats to its status as an important societal institution becomes highly relevant. Such analyses will not eliminate all threats but they might enable educators to re-examine the resources which are now available to them and affirm their present use or plan to redeploy these resources in a more effective fashion.

Changes in the societal context first reveal themselves as trends, but the early detection of trends is a difficult task. In the very early stages a trend is that which is agreed to be so. Once agreement has been reached that a trend exists, the stance which educators may take in response to an identified trend can vary from acquiescence to assertiveness. Whatever the stance taken it is more likely to be effective if

it is taken deliberately. A deliberate position is dependent upon a trend having reached an acceptable level of agreement and clarity.

Methodology

An important question is the degree to which teacher education programs are keeping pace with—or perhaps staying ahead of—social and educational change. To initially explore this question it was first necessary to identify some of the now apparent trends and assess their impact upon the teacher role. Senior undergraduate students then considered how their preparation programs were attending to these emerging needs.

A recent study undertook to gain information upon trends affecting education from two groups. In a preliminary questionnaire study in the Spring of 1983, the purpose was to ask practising educators in Alberta to assist in identifying (1) societal trends, (2) educational trends, and (3) matters impacting on the role of the teacher (Bumbarger & Nixon, 1983). This identification of trends by practitioners was seen as, first, an important element in attempting to overcome the attitudinal separation which exists between educational researchers and educational practitioners (Foster & Nixon, 1978; Levin, 1982) and second, a starting place from which to seek the views of prospective teachers with respect to the existence of these trends and their perceived ability to accommodate to them. Figure 1 is organized under three headings and lists the changing factors that were identified by Alberta educators.

Of primary concern to the teacher educator is the degree to which present preparation programs are a preparation for the dynamic conditions summarized in Figure 1. Ideally programs should be kept in accord with both prevailing societal norms and changing expectations.

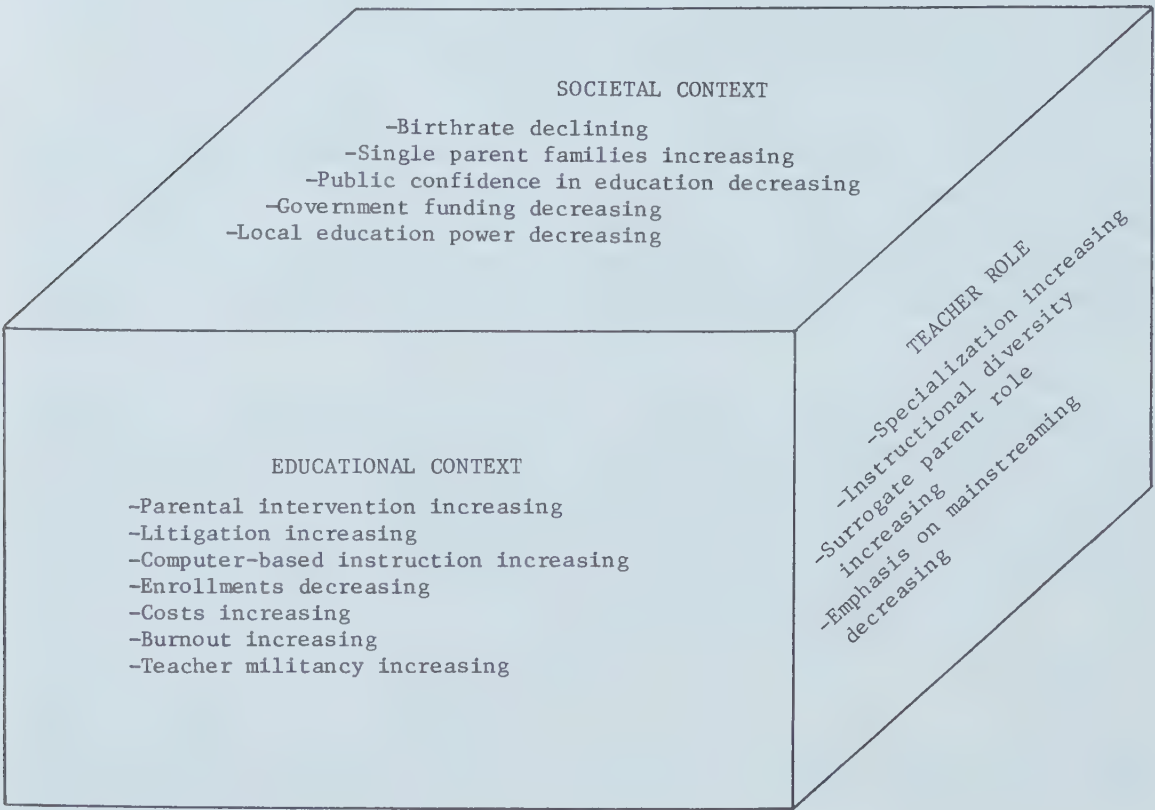


Figure 1. Teacher 'Reading' of the Future

Recognizing that the basic preparation for a teacher is at least four years in length there is clearly the need to be prepared to change the development process even while this basic pre-service preparation is undertaken. Change within courses and change which adds or deletes course work may be necessary but may not be sufficient if the change is needed at the program level.

In a second study at the University of Alberta in January, 1984 senior undergraduates were asked: (1) to give their opinions with respect to the identified trends in the societal and educational context which they saw as increasing/decreasing in impact over the next few years, and (2) to respond in terms of their professional programs as to how well they were being prepared for the identified changes in the role of the teacher (Figure 1).

Data Collection

The data for the educator study were collected as an integral part of a meeting entitled "Role of the Teacher of the Future Conference," held in Edmonton, Alberta, May 12 to 14, 1983. The questionnaire design was developed to provide a pre-conference measure matched with an immediate post-conference measure enabling comparison of pre- and post-test scores for each individual. For this paper pre-conference responses of all registrants (280) were utilized.

Data for the prospective teacher study were collected by means of questionnaires distributed to all students enrolled in the Winter Term, 1984 in the class Educational Administration 401 at the University of Alberta. Of the 530 questionnaires which were distributed a total of 476 completed questionnaires were returned.

Findings and Discussion

Societal Context

Using a Likert-type (5-1) scale, students were asked to respond to the same items asked of practising educators: "Some societal trends may already be discernible. Respond to the following items in terms of where events now happening in society seem to be heading." In Table 1 are the statements of societal trends and the responses of educators and students who thought they would increase/decrease.

TABLE 1

PERCENTAGE OF EDUCATORS AND STUDENTS WHO THINK SOCIETAL TRENDS WILL INCREASE/DECREASE

|                                   | Educators |      |      | Students |      |      |
|-----------------------------------|-----------|------|------|----------|------|------|
|                                   | N         | I    | D    | N        | I    | D    |
| The birthrate                     | 257       | 29.6 | 45.5 | 473      | 28.8 | 41.0 |
| Number of single parent families  | 265       | 86.8 | 5.7  | 476      | 90.8 | 2.5  |
| Public confidence in education    | 266       | 16.5 | 44.0 | 476      | 24.2 | 26.3 |
| Government support for education  | 265       | 11.7 | 52.8 | 476      | 25.8 | 31.7 |
| Local decision power in education | 266       | 31.6 | 45.9 | 476      | 42.0 | 20.4 |



(Increase is indicated by the letter I and decrease by the letter D.) The percentage of respondents who indicated “no change” is not shown in the table. About 93 percent of students could detect change with respect to single parent families but only about half (46 percent) could foresee any change in public confidence in education.

When the educator group mean score responses for these items were compared with the mean scores of students, responses to three items were significantly different: public confidence in education, government support for education, and local decision power in education (Table 2). (Note that these depend more upon opinion than do the other two items.) Educators, perhaps because of increased age and/or increased experience, were more pessimistic in their views than the prospective teachers. However, when the students were asked, “As a prospective teacher, do you think you are being adequately prepared to cope with changes in the societal context as you see them?” less than half (42 percent) felt adequately prepared.

TABLE 2  
COMPARISON OF MEAN SCORES OF EDUCATORS  
AND STUDENTS ON SOCIETAL TRENDS

|                                      | N   | Educator<br>mean | s.d. | N   | Student<br>mean | s.d. | t-value | Probability<br>of t |
|--------------------------------------|-----|------------------|------|-----|-----------------|------|---------|---------------------|
| The birthrate                        | 257 | 2.76             | 1.22 | 473 | 2.84            | 0.96 | -0.97   | 0.334               |
| Number of single<br>parent families  | 265 | 4.15             | 0.80 | 476 | 4.20            | 0.68 | -0.90   | 0.368               |
| Public confidence<br>in education    | 266 | 2.67             | 0.92 | 476 | 2.97            | 0.81 | -4.40   | 0.000 **            |
| Government support<br>for education  | 265 | 2.53             | 0.84 | 476 | 2.93            | 0.87 | -6.06   | 0.000 **            |
| Local decision power<br>in education | 266 | 2.83             | 1.02 | 476 | 3.25            | 0.87 | -5.93   | 0.000 **            |

\*\* significant at the .01 level  
Likert-type scale 5 - 1; 5 indicated Increase and 1 Decrease

Some of the students who considered themselves “adequately prepared” attributed this to the program (e.g., course work in sociology, psychology) but another view was that the “ability to cope” was self-determined.

I think whether we can cope with future change in the societal context depends most on our individual attitudes toward change.  
Not by the program—perhaps on our own.

Of those students who felt less prepared to adapt to societal change, one expressed view was on the limitations of any program.

No school can adequately prepare a person for something that they will only learn how to handle with experience.  
... you can never be fully prepared. One must learn something as one goes.

A more common tendency, however, was to assign “blame” for lack of preparation and to suggest one or more remedies.

We are not being prepared for many of the changes which will occur in the near future. We need more instruction in new programs—reading, computers, math.  
We have not had sufficient instruction in social/economic/political influences on education. Nor have we had sufficient dialogue in problems and suggested solutions.

Educational Context

The procedures used to elicit views on societal trends were again used to seek the perceptions of participants with respect to educational trends. The statements of educational trends and the responses of educators and students who thought that they would increase/decrease are given in Table 3.

TABLE 3  
PERCENTAGE OF EDUCATORS AND STUDENTS WHO THINK  
EDUCATIONAL TRENDS WILL INCREASE/DECREASE

|                            | Educators |      |      | Students |      |      |
|----------------------------|-----------|------|------|----------|------|------|
|                            | N         | I    | D    | N        | I    | D    |
| Parental intervention      | 266       | 72.9 | 8.6  | 476      | 72.3 | 6.7  |
| Recourse to litigation     | 262       | 67.6 | 3.1  | 456      | 45.4 | 3.1  |
| Computer-based instruction | 267       | 92.5 | 0.7  | 476      | 95.6 | 0.6  |
| Enrolments                 | 266       | 27.1 | 41.0 | 471      | 32.7 | 30.8 |
| Costs                      | 267       | 91.4 | 3.0  | 474      | 93.0 | 0.8  |
| Teacher burnout            | 265       | 79.6 | 2.3  | 474      | 68.8 | 4.4  |
| Teacher militancy          | 265       | 61.9 | 9.4  | 471      | 42.7 | 16.6 |

More than 90 percent of students predicted that computer-based instruction would increase as would education costs; 72 percent foresaw an increase in parental intervention; and 69 percent thought that the incidence of teacher burnout would increase. On five of the educational trends the responses given by students were significantly different from those given by educators: (1) recourse to litigation, (2) computer-based instruction, (3) enrolments, (4) teacher burnout, and (5) teacher militancy (Table 4).

TABLE 4  
COMPARISONS OF MEAN SCORES OF EDUCATORS  
AND STUDENTS ON EDUCATIONAL TRENDS

|                            | Educator |      |      | Student |      |      | t-value | Probability<br>of t |
|----------------------------|----------|------|------|---------|------|------|---------|---------------------|
|                            | N        | mean | s.d. | N       | mean | s.d. |         |                     |
| Parental intervention      | 266      | 3.78 | 0.83 | 476     | 3.75 | 0.73 | 0.49    | 0.624               |
| Recourse to litigation     | 262      | 3.79 | 0.75 | 456     | 3.47 | 0.64 | 5.77    | 0.000**             |
| Computer-based instruction | 267      | 4.36 | 0.66 | 476     | 4.49 | 0.60 | -2.64   | 0.009**             |
| Enrolments                 | 266      | 2.86 | 1.04 | 471     | 3.07 | 0.94 | -2.75   | 0.006**             |
| Costs                      | 267      | 4.30 | 0.74 | 474     | 4.32 | 0.63 | -0.38   | 0.701               |
| Teacher burnout            | 265      | 4.15 | 0.81 | 474     | 3.86 | 0.82 | 4.64    | 0.000**             |
| Teacher militancy          | 265      | 3.71 | 0.92 | 471     | 3.34 | 0.91 | 5.41    | 0.000**             |

\*\* significant at the .01 level  
Likert-type scale 5 - 1; 5 indicated Increase and 1 Decrease

Students were more likely than educators to foresee increases in computer-based instruction and education costs, but educators were more likely than students to foresee increases in teacher burnout, teacher militancy and recourse to litigation. Again these differences might be explained by students' inexperience of the "real world" of teaching and/or the optimism of youth.

When students were asked if their preparation programs were adequately preparing them for these changes in the educational context there was less pessimism than was the case with respect to changes in the societal context. More than half the students (54 percent) thought that they were receiving adequate preparation in order to cope with forthcoming changes in education. Reflective of students' comments who saw themselves as adequately prepared were the following:

Many changes in the educational context are brought to the prospective teacher's attention.

We are introduced to new teaching programs at university.

I am better prepared for educational change than societal change.

Preparation more adequately deals with educational concerns.

However, for the sizable minority who felt less satisfied with their teacher preparation there were many concerns but the perceived lack of computer literacy was the one concern which received the highest number of comments. Some students were angered by the lack of computer courses available to them and many made the case for compulsory computer courses.

I resent the fact that I will have to know how to use computers to teach; yet, the university does not deal with it.

We are not prepared to work with computers. A computer course should be compulsory.

Few courses which fit into my very "prescribed" program deal with technological advances, e.g., computers.

Other concerns which were voiced included mainstreaming, "outdated" course content and irrelevancy of some aspects of the program, but no other respondent was as badly off as the student who wrote, "I am so unprepared, I don't even know what you mean by educational context."

### *Impact on the Teacher Role*

Both practising educators and prospective teachers were asked whether they agreed/disagreed that a number of identified factors in the environment would make an impact on the role of the teacher. A 5 point Likert-type scale was again employed with 5 or 4 being used to indicate agreement and 2 or 1 to indicate disagreement. The identified factors and the responses of educators and students who agreed/disagreed that they would have an impact on the role of the teacher are presented in Table 5. Sixty percent of students agreed that teachers would increasingly take on a surrogate parent role but there was no marked apprehension that the other factors would make any major impact on the role of the teacher; subject specialists would still be required and the emphasis on a variety of instructional modes would not decrease. When the mean scores of students were compared with those of educators there were some differences. Educators were more inclined than students to emphasize both the role of the teacher as a surrogate parent and the continued need for diverse instructional modes (Table 6).



TABLE 5  
PERCENTAGE OF EDUCATORS AND STUDENTS WHO  
AGREE/DISAGREE ON THE FUTURE IMPACT ON ROLE OF THE TEACHER

|  | Educators |      |      | Students |      |      |
|--|-----------|------|------|----------|------|------|
|  | N         | A    | D    | N        | A    | D    |
| Technological advances will reduce the need for subject specialists              | 267       | 16.5 | 73.8 | 475      | 14.3 | 67.4 |
| Decreased emphasis on diverse modes of instruction                               | 261       | 12.3 | 69.3 | 472      | 12.5 | 60.4 |
| Teachers will increasingly find it necessary to act as surrogate parents         | 265       | 71.3 | 11.3 | 476      | 58.8 | 14.1 |
| Concerns for efficiency will override the social goals implicit in mainstreaming | 263       | 44.9 | 21.3 | 474      | 37.3 | 25.7 |

TABLE 6  
COMPARISON OF MEAN SCORES OF  
EDUCATORS AND STUDENTS: IMPACT ON TEACHER ROLE

|  | Educator |      |      | Student |      |      | t-value | Probability of t |
|--|----------|------|------|---------|------|------|---------|------------------|
|  | N        | mean | s.d. | N       | mean | s.d. |         |                  |
| Technological advances will reduce the need for subject specialists              | 267      | 2.17 | 1.11 | 475     | 2.27 | 0.99 | -1.26   | 0.208            |
| Decreased emphasis on diverse modes of instruction                               | 261      | 2.22 | 1.03 | 472     | 2.38 | 0.92 | -1.99   | 0.047*           |
| Teachers will increasingly find it necessary to act as surrogate parents         | 265      | 3.80 | 0.98 | 476     | 3.57 | 0.93 | 3.29    | 0.001**          |
| Concerns for efficiency will override the social goals implicit in mainstreaming | 263      | 3.29 | 0.98 | 474     | 3.14 | 0.91 | 2.09    | 0.037*           |

\* significant at the .05 level  
\*\* significant at the .01 level

Likert-type scale 5 -1; 5 indicated Strongly Agree and 1 Strongly Disagree

Anticipating the Future

By and large students did not suffer from inadequacies of self-image. In comparison with their peers they were confident in their personal abilities but they were less confident in the abilities of teacher educators to provide them with programs which enhanced these abilities (Table 7). On a 5-point scale students were asked to rate their abilities/personal attributes as future educators “in comparison to others in the teacher education program.” Fifty-two students (11 percent) rated themselves “outstanding” and another 275 students (60 percent) considered themselves “above average.” Teacher education programs did not fare as well; 5 students

(1 percent) considered their programs “excellent” and 134 students (29 percent) rated them “above average.”

TABLE 7

STUDENTS' SELF RATINGS OF ABILITY  
AND RATINGS OF TEACHER PREPARATION PROGRAM

|             |             | Excellent |     |     |      | Adequate |      |    |      | Poor |      | Total |       |
|-------------|-------------|-----------|-----|-----|------|----------|------|----|------|------|------|-------|-------|
|             |             | 5         |     | 4   |      | 3        |      | 2  |      | 1    |      |       |       |
| Self Rating |             | N         | %   | N   | %    | N        | %    | N  | %    | N    | %    | N     | %     |
| 1           | Outstanding | 2         | 3.8 | 19  | 36.5 | 19       | 36.5 | 5  | 9.6  | 7    | 13.5 | 52    | 100.0 |
| 2           |             | 1         | 0.4 | 86  | 31.3 | 129      | 46.9 | 53 | 19.3 | 6    | 2.2  | 275   | 100.0 |
| 3           | Average     | 2         | 1.9 | 24  | 22.2 | 52       | 48.1 | 29 | 26.9 | 1    | 0.9  | 108   | 100.0 |
| 4           |             | 0         | 0.0 | 5   | 20.8 | 11       | 45.8 | 6  | 25.0 | 2    | 8.3  | 24    | 100.0 |
| 5           | Poor        | 0         | 0.0 | 0   | 0.0  | 0        | 0.0  | 0  | 0.0  | 0    | 0.0  | 0     | 0.0   |
| Total       |             | 5         | 1.1 | 134 | 29.2 | 211      | 46.0 | 93 | 20.3 | 16   | 3.5  | 459   | 100.0 |

The concerns of many students with respect to their preparation programs were of a practical nature. Certain educational trends such as increased computer based instruction in the schools have become apparent but many students felt that they were less competent in this area than they would like to be. Of the 367 students who felt a need beyond the program requirements for further course work, most chose to elaborate on their responses. While courses were sought by individuals on just about every aspect of teacher preparation—from classroom management to knowledge of the content in disciplines related to their majors/minors—84 (almost one-quarter) specified a need for computer courses. In a much smaller number of cases it was thought that more courses might help dispel feelings of apprehension that they were unprepared in a general way for their future as educators.

- I want to take as many good courses as I can.
- I do not, for some reason, feel very confident about teaching. Maybe a few more courses would give me more information so creating more confidence.
- I have not received what I would call a liberal education. I am not educated and therefore not qualified to teach.

Student responses to the question, “When you have completed your professional program and (with a little good luck) have obtained the position of your choice, what aspects of your present program will probably prove most useful to you?” have been summarized in Table 8. Over 90 percent considered the practicum valuable but there was much less agreement with respect to the other program components. Fifty-six percent of students understood the relevancy of curriculum and instruction courses even though these were not always popular. As one student wrote:

- As much as I hated the C. & I. courses and as much as I thought they could have been a lot better, I think they were still useful.
- Education courses (other than curriculum and instruction) and courses outside of the Faculty were much more open to question. The following view was a rare one:  
More emphasis should be put on subject material. Background is necessary to instruction.

TABLE 8  
ASPECTS OF PROGRAM CONSIDERED RELEVANT TO FIRST POSITION

|   | B.Ed. (Elem)<br>(N=325) |      | B.Ed. (Sec)<br>(N=112) |      | B.Ed.<br>(Ind/Voc)<br>(N=32) |      | Total<br>(N=469) |      |
|---|-------------------------|------|------------------------|------|------------------------------|------|------------------|------|
|   | N                       | %    | N                      | %    | N                            | %    | N                | %    |
| Practicum   | 314                     | 96.6 | 96                     | 85.7 | 23                           | 71.9 | 433              | 92.3 |
| C & I courses   | 198                     | 60.9 | 55                     | 49.1 | 9                            | 28.1 | 262              | 55.9 |
| Other education<br>courses  | 110                     | 33.8 | 19                     | 17.0 | 11                           | 34.4 | 140              | 29.9 |
| Courses outside the<br>Faculty of Education<br>related to major/minor | 65                      | 20.0 | 64                     | 57.1 | 16                           | 50.0 | 145              | 30.9 |
| Other   | 28                      | 8.6  | 8                      | 7.1  | 2                            | 6.3  | 38               | 8.1  |

Includes 22 B.Ed. (Elem) After Degree Students  
Includes 14 B.Ed. (Sec) After Degree Students  
Includes 2 B.Ed. (Ind/Voc) After Degree Students

The view of the student who wrote, “You need the theory but experience is much more relevant” was the most prevalent and was expressed in a variety of ways.

We need more practicum!

More CI courses and more practicum.

I think more time should be spent on practical experiences. I learn a lot more through practical experiences than reading books.

I don’t feel I was fully prepared for life in the classroom from courses. Much information was gathered during my practicum.

Most students agreed that there will be a need to update their preparation on a continuing basis (Table 9) but there is some evidence that they feel that their present preparation could be improved. However, the emphasis which students would place on the overriding importance of the practicum may well be misplaced. Overemphasis on the practicum will not only be a step into the past (Dickie, 1983; Patterson, 1982) with its attendant temptation to internalize “by practice” someone else’s program/curriculum/mythologies, but this approach chains one to a static view of teaching. Stated differently, unless there now exists an ideal-teacher model which can be presented for emulation, thoughtful and creative change efforts have to be applied to current formulations. Given that such a model did exist, preoccupation with practicum experiences would still be a vote for the status quo.

A teacher preparation program is made up of a number of course concentrations. Universities usually rely on the judgement of those individual instructors who offer these courses to define the content and to be aware of its relevance to the present and future needs of students. In turn, instructors rely on the feedback from students with respect to the adequacy/relevancy of the material presented. This may not be a sufficient data source but it is often the only data source utilized since courses are evaluated and modified much more often than are entire programs.



TABLE 9  
STUDENTS' SELF-REPORTED NEED BY PROGRAM  
FOR PROFESSIONAL UPDATING

|       |                      | B.Ed. (Elem) |       | B.Ed. (Sec) |       | B.Ed. (Ind/Voc) |       | Total |       |
|-------|----------------------|--------------|-------|-------------|-------|-----------------|-------|-------|-------|
|       |                      | N            | %     | N           | %     | N               | %     | N     | %     |
| 5     | Essential            | 226          | 69.5  | 74          | 66.7  | 23              | 71.9  | 323   | 69.0  |
| 4     |                      | 82           | 25.2  | 26          | 23.4  | 9               | 28.1  | 117   | 25.0  |
| 3     |                      | 14           | 4.3   | 7           | 6.3   | 0               | 0.0   | 21    | 4.5   |
| 2     |                      | 3            | 0.9   | 3           | 2.7   | 0               | 0.0   | 6     | 1.3   |
| 1     | Not essential at all | 0            | 0.0   | 1           | 0.9   | 0               | 0.0   | 1     | 0.2   |
| Total |                      | 325          | 100.0 | 111         | 100.0 | 32              | 100.0 | 168   | 100.0 |

Includes 22 B.Ed. (Elem) After Degree Students  
Includes 14 B.Ed. (Sec) After Degree Students  
Includes 2 B.Ed. (Ind/Voc) After Degree Students

Students are one data source in the assessment of programs but other sources which should be considered include graduates of particular programs, school district personnel, and teachers' associations. Graduates may well view their preparation programs differently than do students. For example, it is possible that over time the importance attributed to the practicum experience decreases and the relevance of the other program components to the practitioner becomes clearer. District personnel with responsibilities for hiring may not be able to evaluate the differences among teacher education programs but they are likely to be able to offer projections with respect to the need for teachers with particular specializations. Professional associations can monitor teachers' needs for in-service and professional renewal. If Faculties of Education are to provide programs which are forward looking, the co-operative nature of the teacher education task must be understood (Nixon & Bumbarger, 1983).

There has been much written and said about change and its inevitability as well as its increasing pace. Perhaps it is time to reconceptualize what we mean by change—and adaptations to it that can be made. A major restructuring of teacher education programs based on some vision of the role of the teacher in the twenty-first century is neither feasible nor necessary. Students in education programs are future teachers but their preparation is now. All that can reasonably be expected is that we provide students with skills and knowledge to meet presently identifiable changes in the role of the teacher. If this argument is accepted, then the future is now and the responsibility to change to meet it is now.

References

Apple, M. W. (1983). Curriculum in the year 2000: Tensions and possibilities. *Phi Delta Kappan*, 64(5), 321-326.

- Bakalis, M. J. (1983). Power and purpose in American education. *Phi Delta Kappan*, 65(1), 7-18.
- Bernstein, H. T. (1983). The Information Society: Byting the hand that feeds you. *Phi Delta Kappan*, 65(2), 108-112.
- Bumbarger, C. S., & Nixon, M. (1983). Alberta Teachers' Association Conference Role of the Teacher of the Future. *Measuring participants' opinions*. Department of Educational Administration, The University of Alberta.
- Dickie, B. T. (1983). *Portrait of a practicum*. Unpublished PhD thesis, University of Alberta.
- Foster, L., & Nixon, M. (1978). Teacher as researcher: The problem reconsidered. *Canadian Journal of Education*, 3(1), 75-88.
- Gideonse, H. D. (1982). The necessary revolution in teacher education. *Phi Delta Kappan*, 64(1), 15-18.
- Gold, S. (1983). Victims of the economic crisis. *Alberta Teachers' Association Magazine*, 63(4), 8-11.
- Goodlad, J. I. (1982). Let's get on with the reconstruction. *Phi Delta Kappan*, 64(1), 19-20.
- Levin, B. (1982). Research and practice: The place of the school district research unit. *Ninth yearbook 1982 the study of education: Canada, 1982*. Canadian Society for the Study of Education, 9-15.
- Moore, W. E. (1963). *Social change*. Englewood Cliffs, NJ: Prentice-Hall.
- Nixon, M., & Bumbarger, C. S. (1983). Tomorrow's teachers: A shared responsibility. *The Saskatchewan Educational Administrator*, 15(2), 31-41.
- Patterson, R. S. (1982). Alberta's normal schools. *Alberta Teachers' Association Magazine*, 62(2), 8-13.
- Ravitch, D. (1983). On thinking about the future. *Phi Delta Kappan*, 64(5), 317-320.
- Toffler, A. (1981). *The third wave*. New York: Bantam Books.

*Errata:* The publication of the paper, *Recollections about the life of the first Ukrainian settlers in Canada*, by W. A. Czumer, listed on page 86 of the March, 1984 issue of the journal was incorrectly attributed to the University of Toronto. The paper was published by The Canadian Institute of Ukrainian Studies, University of Alberta, Edmonton, Alberta.

# CONTEMPORARY EDUCATION

Now in its fifty-fifth year, **Contemporary Education** continues to reach all segments of the educational community: learners, teachers, professors, professional educators employed by local, state, and federal agencies, and administrators. Some recent articles illustrate the wide range of topics covered:

Violations of Management Principles Within Academe

Andrew F. Sikula and John P. Sikula

Why Not Merit Pay?

Lawrence J. Dennis

Church, State, and Moral Education

Richard LaBrecque

The Curriculum of Capitalism

Mark Herring

The Promotion Game

William G. Chance and Sandra Bone

Psychology of Learning: Theory and Application

P. G. Aaron and Ebrahim Fakouri

**Contemporary Education** seeks to present competent discussions of contemporary problems in education and generally restricts its contributing personnel to those of preparation and experience in the field. Unsolicited manuscripts are welcome.

The annual subscription rate is \$9.00 within the United States and \$12.00 to foreign countries including Canada and Mexico. Back issues are available at \$3.00 per copy.

Address all communications to:

Russell L. Hamm, Editor  
Contemporary Education  
1005 School of Education  
Indiana State University  
Terre Haute, Indiana 47809



# THE JOURNAL OF EDUCATIONAL THOUGHT

Published three times yearly by the Faculty of Education  
The University of Calgary

Telephone (403) 284-5629

The Journal of Educational Thought promotes speculative, critical and historical research concerning the theory and practice of education in a variety of areas: administration, comparative education, curriculum, educational communications, evaluation, instructional methodology, intercultural education, philosophy, psychology, and sociology. Critical and controversial views on the field of education or on previously published articles in The Journal of Educational Thought appear in a section entitled The Forum, which also contains papers of a less formal nature. The book section includes short book reviews as well as extended book review essays. The Journal is international in scope, qualitative in nature, and serves a broad readership; specialists in the areas mentioned, scholars, and the public in general.

---

**Instructions for the preparation of manuscripts appear on the last page of each issue.**

---

**Publishers are invited to submit to the Editors books for review.**

---

The Journal is indexed in *Canadian Education Index* and by the *Educational Resources Information Centre (ERIC)*. Abstracts of appropriate articles appear in *Educational Administration Abstracts*, *Language and Language Behaviour Abstracts*, and *Sociology of Education Abstracts*. Available on microfilm from University Microfilms, Inc., Ann Arbor, Michigan.

It is the policy of the Journal to provide an open forum for scholarly discussion. Editorials as well as articles are to be regarded as expressing the viewpoints of the writers themselves, and not as reflecting official positions of the Editorial Committee, the Faculty of Education, or The University of Calgary. It is an additional policy of *THE JOURNAL OF EDUCATIONAL THOUGHT* to modify at its discretion wordy and sexist expressions appearing in a manuscript submitted for publication, whenever a change in these expressions will not distort the meaning of the passage.

---

## Subscription Rate

Effective January 1st, 1982

\$15.00 - 1 year (individual)  
\$20.00 - 1 year (institutions)  
\$ 6.00 - single issue  
\$ 2.00 - back issue  
\$ 7.00 - Special issue

## Address

The Journal of Educational Thought  
Faculty of Education  
The University of Calgary  
2500 University Drive N.W.  
Calgary, Alberta, Canada  
T2N 1N4

Postage overseas air printed (not U.S.), additional \$2.00.

**Claims for undelivered copies must be made within one month following the month of publication.**

# New Zealand Journal of Educational Studies

---

Published twice-yearly by the New Zealand Council for Educational Research

The New Zealand Journal of Educational Studies provides an avenue of publication for essays, research findings and critical comments in any of the major fields of educational study in New Zealand and the Pacific area.

The current subscription rate is US \$12.00

Business correspondence and subscriptions should be addressed to the Executive Officer, New Zealand Council for Educational Research, Box 3237, Wellington, New Zealand.

## New Zealand Journal of Educational Studies

Some recent articles:

---

Performance Feedback on Written Tasks for Low-achieving Secondary Students

*J. Scriven and E.L. Glynn*

Pre-school Screening for Learning Disabilities: Is it Warranted?

*Christina van Kraayenoord*

Teacher Turnover in an Economically Depressed Area

*P.D.K. Ramsay and D. Sneddon*

Corporal Punishment and Attitudes to Violence of Secondary School Students

*Jane Ritchie*

Curriculum Development in a Cross-Cultural Setting: Social Science in the South Pacific

*J.M. Renner*







## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by authors' last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary* or to the *Dictionary of Canadian English*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.





# *ajer*

THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH

VOLUME XXX

NUMBER 4

DECEMBER 1984

PUBLISHED BY  
THE UNIVERSITY OF ALBERTA • EDMONTON

# THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

Published in March, June, September, and December by the  
Faculty of Education, The University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

EDITOR: Walter H. Worth

## CONSULTING EDITORS

D. C. Bjorkquist  
University of Minnesota

P. E. Bryant  
University of Oxford

B. M. Chung  
Han Rim University,  
Republic of Korea

A. Côté  
University of Ottawa

D. Fitzgerald  
The University of New England,  
Australia

D. B. Harrison  
The University of Calgary

N. L. Hersom  
The University of Saskatchewan

M. Laferrière  
McGill University

T. Morrison  
University of Manitoba

J. Richardson  
University of New Brunswick

P. Robinson  
Atlantic Institute of Education

T. Rogers  
University of British Columbia

R. E. Traub  
Ontario Institute for  
Studies in Education

K. Watson  
University of Reading

R. J. Wilson  
Queen's University

*The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Alberta Education Index*, *Canadian Education Index*, *Current Contents/Social and Behavioral Sciences*, *ERIC/Current Index to Journals in Education*, *Multidisciplinary Index to the World's Social Sciences Literature*, and *Social Science Citation Index*; appropriate articles are abstracted in *Cognitive Development Abstracts*, *Educational Administration Abstracts*, *Educational Technology Abstracts*, *Language Behavior Abstracts*, *Multicultural Education Abstracts*, *Psychological Abstracts*, *Research into Higher Education Abstracts*, *School Organization and Management Abstracts*, and *Sociology of Education Abstracts*, and reviewed by the *Journal for Research in Mathematics Education*.

The subscription rate is \$18.00 per year for individuals, \$24.00 per year for institutions. Add \$6.00 for delivery outside of Canada. Single copies are \$6.00 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, The University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.



# The Alberta Journal of Educational Research

Volume XXX, Number 4

December, 1984

## CONTENTS

|  |     |
|--|-----|
| Teachers' Perspectives on Inservice Education . . . . .  | 239 |
| <i>M. Pansegrau</i>  |     |
| One Last Word List . . . . .   | 259 |
| <i>L. Gunderson</i>  |     |
| Communication Skills in Immersion Programs . . . . .   | 270 |
| <i>S. T. Carey and J. Cummins</i>  |     |
| Student Teachers' Perceptions of Their Competence Versus<br>the Perceptions of University Supervisors and Sponsor Teachers . . . . . | 284 |
| <i>C. F. Staab</i>   |     |
| An Evaluation of the Diagnostic Efficiency of the<br>Wechsler Intelligence Scale for Children-Revised . . . . .                      | 299 |
| <i>H. H. Mueller, G. J. Mancini and R. H. Short</i>  |     |
| Computer-Assisted Instruction in College Composition . . . . .   | 311 |
| <i>P. Evans, N. Mickelson and N. C. Smith</i>  |     |
| Delegation: The Neglected Aspect of Participation in Decision Making . . . . .   | 320 |
| <i>W. K. Hoy and D. A. Sousa</i>   |     |

## PERSPECTIVES

|  |     |
|--|-----|
| The Theory-Practice Relationship in Teacher Education:<br>Need There Be Any? . . . . . | 332 |
| <i>P. J. Miller and G. D. Taylor</i>   |     |

|   |     |
|---|-----|
| TABLE OF CONTENTS, VOLUME XXX . . . . . | 340 |
|---|-----|

FACULTY OF EDUCATION  
*The University of Alberta*





MORAG PANSEGRAU

University of Alberta

## Teachers' Perspectives on Inservice Education

*The purpose of this study was to determine teachers' perspectives on inservice education. It was postulated that the beliefs and activities—the who, what, when, where and why—of those who participate in, or are expected to participate in, inservice education would fall into patterns and that the portrayal of such patterns would provide some insight into the meaning held by teachers of inservice education.*

*Data were collected over a twelve-month period using the field research strategies of semi-structured, in-depth interviews and participant observation. The findings of the study reveal that teachers classify inservice education into the following four discrete groups of learning activities, each of which they perceive differently—Group One: mandatory, formal inservice education activities offered in conjunction with a compulsory program change; Group Two: other types of mandatory, formal inservice education activities; Group Three: formal inservice education activities of a voluntary nature; and, Group Four: non-formal inservice education activities. Based on the findings of the study, recommendations for school systems and universities are given.*

### *Introduction and Need for the Study*

It has long been recognized in English-speaking countries that preservice teacher education represents but the minimum prerequisite for entry into the teaching profession. Teachers must continue to develop professionally throughout their teaching career. Consequently, inservice education activities have been offered since the late 19th century (Richey, 1957; Stephens & Hartman, 1978; Bacon, 1980). These activities are seen as a means of ensuring that practicing teachers are exposed to new techniques, programs and other innovations and are given assistance with classroom-related and subject-related areas of concern.

The current knowledge explosion and rapid rate of technological growth, the aging and relatively static teacher population, the current public dissatisfaction

---

Morag Pansegrau holds a PhD in Educational Administration and, at the time this article was submitted, was Coordinator of the Committee for Improvement of Teaching and Learning, University of Alberta. She is now Senior Instructor in Business Education, Alberta Vocational Centre, Edmonton, Alberta.

with the products of the educational system, the ever-expanding role being assigned by society to educational institutions and the rapidly changing lifestyle of many of the segments of our society, all contribute to the necessity of ensuring that teachers are involved in an ongoing program of professional development. As Bacon (1980) points out:

even the most excellent preservice program is only a preparation for entry into the profession, and cannot last a whole career because the context teachers work in changes. (p. 1)

Indeed, according to Houle (1980), a zest for knowledge and lifelong learning are key traits that qualify an individual for membership in a profession. Many professions have come to the realization that their members must individually and collectively accept the obligation to continue to learn and that programs to achieve this purpose must be established (Harrington, 1979; Frandson, 1980; Houle, 1980).

Unfortunately, as much of the literature over the past two decades acknowledges, inservice education has been relatively ineffective as a vehicle for introducing educational reform into classrooms. Despite vast expenditures by government, by school boards, by professional councils, by universities and by teachers (Fennell, Hill, & Thiessen, 1980; *ATA News*, 1981), large numbers of teachers appear to view such activities with indifference or hostility, attend reluctantly, if at all, and do not appear to change their classroom behavior in the ways anticipated (Fullan & Pomfret, 1977; Cooper & Hunt, 1978; Brayne, 1980). Consequently, "many of our grand plans for educational change have failed to have major impact or have been co-opted" (West, 1975, p. 35). As Howey and Joyce (1978) point out, inservice education does not appear to be in the best of health, despite Bruce's (1980, p. 140) prediction that, "in the 1980s innovation and growth will be found in the inservice education of practising teachers." Indeed, it has been described as the "slum of American education" (Wood & Thompson, 1980, p. 374). A pertinent question is posed by King, Hayes, and Newman (1977):

Why do teachers, who almost universally appear eager to improve their professional performance, frequently respond with disdain or outright hostility to local efforts to "inservice" them? (p. 686)

In order to combat teachers' inferred indifference to their professional development as indicated by their reluctance to attend inservice education activities organized for them by others, and their unwillingness when they do attend to adopt the information presented, a variety of solutions has appeared in the literature. The solutions currently in vogue include involving teachers in the planning stages, conducting needs assessments and manipulating various externals such as the timing, format and location of activities. What is commonly assumed by those writing about and organizing inservice education is that teachers who attend inservice activities will transfer the learning to the classroom and it will be observed in the changed routine of the classroom. When teachers do not immediately adopt the new behaviors, it is further assumed that the delivery format needs to be altered. The key to successful inservice education thus is seen as the discovery of new and motivating models for the delivery of inservice education, and the literature is replete with criteria, guidelines, models and lists of shortcomings and recommendations, all of which, it is inferred, will result in more effective inservice education.

It does not appear to have been seriously considered that teachers may attend inservice education activities for reasons other than to obtain information which they will use to bring about an immediate change in their classroom behavior, or



that teachers may deliberately choose not to adopt the new behavior. What have not been determined are teachers' perspectives on inservice education.

Perhaps, before suggesting a plethora of antidotes to combat teachers' alleged indifference or hostility towards inservice education, teachers' perspectives on inservice education should be determined. Although there exists a vast array of articles and texts on the topic of inservice education, the field lacks a formal, systematic research base and thus "until inservice teacher education is examined more closely, we can only talk about what we think we know" (Cruickshank, Lorish, & Thompson, 1979, p. 31). In particular, inservice education research has given little attention to descriptive studies (Cruickshank, Lorish, & Thompson, 1979, p. 31). Thus, this study was designed as an inductive, descriptive and analytical study which would permit the researcher to determine how teachers perceive the world of inservice education but would avoid imposing upon teachers the preconceptions of others in relation to teacher inservice education.

Consequently, in this study effort was devoted to an understanding of the total phenomenon of inservice education by determining the beliefs and activities—the who, what, when, where and why—of those who participate in, or are expected to participate in, inservice education activities. It was posited that their beliefs and activities would fall into patterns and that the portrayal of these patterns would provide an insight into the meaning for teachers of inservice education.

### *Methodology and Procedures*

Traditionally, research in the field of education has adopted research methods associated with the natural or physical sciences. As Blumer (1969) states:

the predominant procedure is to take for granted one's premises about the nature of the empirical world and not to examine these premises. . . . Instead of going to the empirical social world in the first and last instances, resort is made instead to a priori theoretical schemes, to sets of unverified concepts, and to canonized protocols of research procedure. (pp. 32-33)

In recent years, however, there has been growing disenchantment with the processes and results of traditional educational research methods and an increasing interest in what, according to Spradley and McCurdy (1972) and Bogdan and Biklen (1982), are generally labeled ethnographic methods. These methods, as West (1977, p. 61) states, "allow one to understand how conceptions held by people shape their behavior, at least in part, and how such conceptions and behavior change over time."

In an exploratory study of a relatively unresearched topic, a methodology and procedures which would permit the researcher to determine what Bruyn (1966) terms "inner perspective," or the understanding of people from their own frame of reference, were sought. The symbolic interactionist perspective (Blumer, 1969, 1978), a methodological stance that respects the nature of the empirical world, guided the research; and, an inductive approach in which the researcher enters the field with an area of interest, some general questions to be answered, but no more, was adopted.

Teachers' perspectives on inservice education were gathered through the field research strategies of in-depth interviewing and participant observation. The major data gathering strategy was a series of arranged in-depth elite and specialized interviews (Dexter, 1970) which used an open-ended, semi-structured approach in order to focus the interview yet permit the interviewee to define the situation, structure his account of it and introduce his notions of what he regarded as rele-

vant. Longhand and shorthand notes were used to record the views of participants and the observations of the researcher.

The researcher had access to teachers and inservice education activities in two school systems. Over a twelve-month period, 77 arranged in-depth interviews were conducted with 86 teachers in 11 schools. In addition, four days of teacher conventions and two professional development days organized for the teachers included in the study were attended. Informal conversations with many teachers and those responsible for organizing inservice education activities for teachers were a part of the data-gathering.

The general approach to inductive research as outlined by Glaser and Strauss (1965, 1966, 1967) and Glaser (1978) prevailed. This permits the researcher to ground the research and theory in the reality he/she is studying and allows "important concepts, basic categories and significant hypotheses" (Glaser & Strauss, 1966, p. 57) to emerge. The data collection process is guided by the principle of theoretical sampling, theoretical saturation and the constant comparative method. The researcher collects, codes and analyzes the data and decides what data next to collect and where to find it; evidence is collected from other comparative groups to check out whether the initial evidence is correct. Theoretical saturation is the criterion for judging when to stop sampling the different groups pertinent to a category; that is, no additional data are being found whereby the properties of the category can be developed. Thus data collection, coding and analysis is ongoing, points the direction for future sampling and determines the focus of future interview probes.

The field was entered with a minimum of predetermined assumptions and a general question to be answered, namely, "What are teachers' perspectives on inservice education?" Although the data collection, coding and analysis steps outlined in this study and often found in the literature may appear to have a linear quality, in reality they often occur simultaneously. Entry may have been under negotiation in one school, interviews underway in another, while data collected from another school were under analysis. The data obtained from the first interviews revealed certain views held by teachers about inservice education and suggested tentative hypotheses to be explored. The first group of interviews was conducted with high school teachers. After the responses were coded and analyzed it was decided to interview teachers at other grade levels to compare their views. The views of teachers in promoted positions were sought to determine if they varied from those of classroom teachers. Interviews were directed to teachers highly receptive to organized inservice activities as well as to those with less enthusiasm for such activities. Data were collected from teachers in elementary, junior high and senior high schools, and from administrators. Both male and female teachers with a variety of subject area specialties and at various stages in their careers were included in the sample. The research also sought to determine if those teachers choosing not to volunteer to participate in the study were radically different from those volunteering; it was judged that there was no discernible difference.

The data collected were analyzed using the coding and category development tactics suggested by Turner (1981) and taking into consideration the suggestions of Glaser and Strauss (1967), Glaser (1978) and Bodgan and Biklen (1982). The four elements of a perspective as outlined by Becker, Geer, Hughes and Strauss (1961, p. 436) guided the organization of the categories and the presentation of patterns. The data were grouped, finally, into 12 categories, and from these categories emerged patterns which reveal teachers' shared perceptions of inservice education (see appendix).



Summary of the Findings

It was the intent of this study to determine teachers' beliefs and activities in relation to inservice education in order to ascertain if there were common patterns which would permit a greater understanding of the phenomenon of teacher inservice education. It was posited that those writing about and organizing teacher inservice education and those attending inservice education activities may be operating from different premises. An analysis of the data collected indicates that teachers share certain perceptions relating to inservice education. These shared perceptions reveal a discrepancy between how teachers perceive inservice education and how others assume teachers to perceive inservice education, assumptions which usually form the rationale for the various inservice education activities organized for teachers.

| VALUE OF AN ACTIVITY  | essential  |   |  |  |   |
|---|--|---|--|--|---|
|   | rewards must equal sacrifices  |   |  |  |   |
|   | worthless  |   |  |  |   |
| GROWTH<br>(higher level needs)<br>↑                                   | to obtain information that will be used to effect a major change in class-room behavior  |   |  |  |   |
|   | to enjoy a mentally stimulating break from routine   |   |  |  |   |
|   | to obtain recognition that a worthwhile job is being performed   |   |  |  |   |
|   | to associate with adults instead of children and exchange information with colleagues  |   |  |  |   |
| NEEDS SATISFACTION<br>(reasons for attending various activities)<br>↓ | to become acquainted with the latest developments in the field of education.   |   |  |  |   |
|   | to obtain confirmation that present practices are appropriate and acceptable   |   |  |  |   |
|   | to obtain information and materials relating to present teaching assignment which can be immediately adopted (requires no major change in ideology or current practices) |   |  |  |   |
| DEFICIENCY<br>( lower level needs)                                    | non-attendance as attending will not satisfy a need  |   |  |  |   |
|   |  | group one: mandatory, formal inservice education activities offered in conjunction with a required program change | group two: other types of mandatory, formal inservice education activities | group three: formal inservice education activities of a voluntary nature | group four: non-formal inservice education activities |

NOMOTHETIC < ----- ORIENTATION ----- > IDIOGRAPHIC

Figure 1. Groups of inservice education activities: A model of inservice education as perceived by teachers



The findings of the study reveal that there is not one single perspective on inservice education held by teachers but rather that teachers perceive inservice education to fall into four discrete groups of learning activities, each of which they perceive differently. Their decision whether or not to participate, their reasons for participating and their assessment of the value of an activity vary according to the particular group of activities in question.

Teachers' shared perceptions of inservice education are portrayed in Figure 1, "A Model of Inservice Education as Perceived by Teachers."

Teachers classify inservice education by types of activities which range from formal activities organized for teachers by others with a nomothetic or institutional orientation, to activities of a non-formal, ongoing nature, that tend to be individualized and teacher initiated, and have an idiographic orientation. Teachers participate in activities of both a formal and non-formal nature in order to satisfy certain needs which range from lower level deficiency needs to higher level growth needs. Although participation in an activity may satisfy more than one need, usually there is a dominant need which the teacher desires to have satisfied and which accounts for his attending one activity rather than another. The value of a particular activity is determined by the individual, his needs at that moment in time, and whether attending the activity satisfied those needs. In general, the activities are regarded as essential, worthless, or worthwhile if the rewards derived from attending equal or exceed the sacrifices or costs that are incurred by attending. Following is a synopsis, by groups of activities, of teachers' perspectives on inservice education, as illustrated in Figure 1.

*Group One: Mandatory, Formal Inservice Education Activities Offered in Conjunction with a Compulsory Program Change*

These activities are usually held during the regular school day and are attended by those who are presently teaching the subject or who will be responsible for this area in the future. Teachers attend formal inservice education activities organized in conjunction with a compulsory program change in order to obtain information which they will use to effect a major change in their classroom behavior (whether such a change actually occurs is outside the scope of this study). They consider these activities to be essential if they are to interpret and implement new programs.

If you are going to use a new curriculum, inservice is an issue, a big deal. Teachers know they are going to have to teach the curriculum and they had better go to the inservice.

In social studies the written prepared units will do the job of telling the teacher what to teach . . . but the one-day inservice will bring it all together. You need both parts to be effective.

They prefer sessions held during the school day (with students released from classes) and they want relevant information and materials presented by classroom teachers with expertise in the area.

I feel much better this year with a day off for social studies inservice, you're not looking at the clock after 4 p.m. . . . It's the first time for a one-day inservice in school time, in depth, done by a teacher not a publisher, and it makes it worthwhile.

It's of benefit having no students on the day of the social studies inservice as then you don't have to prepare for subs, worry about the class clown, tie up all the loose ends on your return to school.

*Group Two: Other Types of Mandatory, Formal Inservice Education Activities*

Included in this category are formal inservice education activities held during the regular school day which teachers are obligated to attend (students are released from classes). Such activities include professional development days, teachers' conventions and sessions organized by the school or school system.

Much of the criticism leveled by teachers against this type of activity seems to have its roots in the fact that there is no long-term plan for continuing teacher professional development. Rather, activities are offered in a haphazard, piecemeal fashion, usually the topics are of a "one of" nature, tend to be repetitious, are offered at only one level and do not cater to the different needs of teachers at various stages in their career.

No attempt is made to match the needs of teachers with what is organized, no attempt at all.

P.d. is offered piecemeal. You are made to go to activities but there's no plan when you compare it to other professions . . . they insist on certain things and perhaps they make people more aware of what they should be looking at (re p.d.). We don't do that in Education yet.

Teacher inservice education needs follow an identifiable pattern which changes over time—a pattern that is consistent with the literature and research on adult learning and development, the nature of professional learning and the concerns of teachers.

There is a pattern, a change. At first you are insecure, you have to do a great job in many areas and you attend inservice randomly, with not much selection. As a beginning teacher I would go to just anything . . . then after I had taught a few years I became more selective.

It appears that at the beginning of their teaching career, teachers feel insecure and overwhelmed and desire formal inservice education activities that will help them cope with their teaching assignments.

When I first started I felt incomplete and I went to everything. Now I feel more competent and I know how to teach reading, I have used four different language arts series . . . so I don't go to language arts sessions.

As they become more experienced, they gain in confidence and knowledge and often prefer more thought-provoking presentations.

I think the beginning teacher tends to look for more things that could help him in the classroom be a better teacher as he is groping. A more experienced teacher feels more confident, goes to more abstract, inspirational things, probably that's why I prefer conventions now.

However, whenever they make a subject or grade change or are promoted, initially they function as neophytes in the new position.

After the first degree, the first few things you go to you want someone to tell you what to do on Monday morning. Once you've got the day-to-day under control you are more theory oriented. . . . Same applies when you make a change.

The teachers included in this study believe that there are many ways in which teachers can continue their professional development and that attending formal inservice activities is but one way.

Over ten years I have seen every approach from the evangelical to the practical, hands-on and now it seems that there is a need on p.d. days to let me come to school in my jeans and have time to myself in my personal school environment.

Teachers say, let me as a professional say where I want to put the time. . . . I really do p.d. activities at other times, not necessarily on official professional development days.



Although attending these activities permits teachers to satisfy certain needs, the dominant reason for attending is not a desire to obtain information which they will use to effect a major change in their classroom behavior.

I don't necessarily go and say, "Hey, I am going to get new ideas for my classroom."

I think it is more of a recharging experience and if you believe what you hear you may change your behavior but I don't think, honestly, that happens.

Consequently, these activities may be considered to be less than successful when judged by the criterion of an overt change in the classroom behavior of teachers. On the other hand, the activities may be considered eminently successful when judged by the various criteria identified by the teachers included in this study. Although to organizers who expect to see an overt behavior change result from teacher participation, the needs identified by teachers may appear to be inconsequential, almost frivolous, and the benefits derived nebulous, to teachers these needs are very important and their satisfaction exceedingly desirable, if not essential, for their mental health and sense of well-being. Indeed, although teachers would like to see changes made, many consider these activities to be worthwhile, regardless of topic or presenter.

You get a great deal of value from organized inservice activities, even if you go to bad ones . . . they're stimulating, even despite the speaker.

Following are the various needs identified by teachers that are satisfied by attending mandatory, formal inservice education activities that do not accompany a compulsory program change. The various needs range from lower level needs of a safety, security and belonging nature to those at a higher level that promote growth.

A. *Non-attendance.* Some teachers choose not to attend these activities, even though they are held during school hours and students are released from classes. These teachers consider school days without students to be essential, but do not approve of the ways in which these days currently are organized. As adults and professionals, they want the freedom to choose if, and when, to participate in formal inservice education activities. They believe that on days when students are released from classes teachers should have the right to engage in whatever types of learning activities they, not others, consider will be most beneficial to them and to their students.

We never say: What would you like to do to be a better teacher?

There should be a choice of activities—do what you want whether it is socializing or workshops.

Teachers go to conferences, workshops, we think are important.

B. *To obtain information and materials relating to present teaching assignment which can be immediately adopted (requires no major change in ideology or current practices).* Teachers are constantly seeking new ideas, techniques, behaviors and materials which they can use to enliven their present classroom practices but which require no fundamental change in ideology or current practices. At these sessions they desire practical, classroom-related information which they can put into practice immediately and which requires only those materials currently available in their schools. They prefer presentations given by individuals who are, or were until recently, classroom teachers and who understand the realities of life in the classroom. Too, they want hands-on activities which permit them to become involved and handouts, as such take-away information greatly enhances the likelihood of their using the information presented.



I prefer highly technical workshops, with teaching techniques, lesson plans—I want concrete activities for the classroom.

I want hands-on inservice using materials that are available in our school. I don't want to be given ideas that you can't carry out yourself.

I want hands-on, small group activities given by presenters who have dealt with the same age level and situation as you . . . you forget what the classroom is like . . . something to pick up and take with you—the only way it will come back to the classroom.

*C. To obtain confirmation that present classroom practices are appropriate and acceptable.* It has been documented that teachers accept the findings of educational research that are neutral or confirm present practices, otherwise findings tend to be ignored (Clifford, 1973, p. 27). The same appears to be true for formal inservice educational presentations.

We believe what is relevant and what backs up our position; we don't remember or distort what is not acceptable to us.

Teachers want to attend positive presentations which enhance their self-concept by reassuring them of the appropriateness of their present practices.

You listen to someone and they tell you to do this and that and you are doing it and you say, "Hey, that's what I'm doing," and that's good. There's a lack of pats on the back for teachers, there's no recognition.

I felt so good when the presenter mentioned certain things I was doing.

They indicated that unless they were in a crisis situation which necessitated an immediate change in their classroom behavior (for example, a compulsory program change), they will not adopt a new approach, even if they consider it to have merit.

Unless it's an area I don't know too much about I go for confirmation, not a conscious desire to change unless I feel shaky in the area. I wouldn't adopt right away unless I felt dissatisfied with my approach, I've not got the time.

*D. To become acquainted with the latest developments in the field of education.* Teachers wish to keep current, and desire presentations which will expose them to new knowledge. Although they admit that they probably will not adopt the new concept, they state that they are open to the seeds of a new idea being sown. However, they want the new knowledge to germinate and come to fruition if and when they decide that the concept is acceptable and the timing appropriate. If they change their behavior, usually it is after several exposures over an extended period of time.

To change behavior takes years, but inservice acquaints teachers, stimulates them and if they choose to follow up, that's their decision.

I get ideas from speakers, and encouragement, but I think it is a slow assimilation.

It may even take more than one inservice on a topic before you consider using it.

Indeed, even if they desire to adopt a new idea, they would wait until "next time around"; there are simply too many daily tasks that must be performed.

Even if you change because of inservice, you probably will not change that year or the next; you might adopt a couple of ideas a couple of years later . . . changing a program is really hard. I try to, but it puts more pressure on you. . . . It's odd, whether you are a first year or a ten year teacher, none of us like the pressure and the pressure does not lessen.

I tend to not change behavior, partly because of lack of time . . . I would adopt at the beginning of the next time I was teaching the unit. It's bad enough just getting through the day-to-day with 101 things to do . . . I've not got two weeks to set up a new program.

It was also suggested that teaching does not attract innovators.

Part of the reason people get into teaching is that they are very conservative, rational people—if they were radical, they would not last, or like, teaching. . . . You tend to go to inservice on, for example, dyslexia, listen, return to the classroom and pick up at the chalkboard where you left off, and make no attempt to use the information on learning/reading disabilities.

Some pointed out that it is easier to bring about change if others in the school are knowledgeable and supportive.

It's difficult to apply new ideas, concepts, if other staff members are not interested.

It was suggested that changing behavior after an inservice which accompanies a compulsory program change is easier because the teacher does not have to convince others of the need to change.

To adopt, you must justify new ideas to bosses, colleagues, incorporate them into your present teaching methods. It's different with social studies inservice as the curriculum is in use and you will be evaluated on it.

However, as several teachers mentioned, those involved must feel the need to change.

There must be some sort of dissatisfaction with what is, and if you go to a workshop and you feel dissatisfied, then you are tempted to change . . . unless you can create that need for change, it is very hard. The problem of inservice is that if you have to inservice people, it must be linked to a need, and you have to create that need for change and if it is not created, the whole inservice is a waste of time.

We wanted to change a couple of teachers in a certain school, make them more empathetic. We gave a session on discipline. The ones already empathetic became more so, felt positive, expanded, but the ones we wanted to change, it was like water off a duck's back—that was us imposing our view of their needs on them and they did not perceive it as a need and did not respond to the session.

In the final analysis:

Whether you use, implement the idea depends on what it is, how it's presented and if you were interested in the first place.

E. *To associate with adults instead of children and exchange information with colleagues.* The isolation of the classroom teacher appears to be a cause of great concern to teachers. Attending organized inservice education activities ensures that teachers obtain a much needed mental health break from the regimented world of the school and the isolation of the classroom.

I like kids . . . but I need to escape to an adult level intellectually.

I think that getting out of the classroom and exchanging ideas is a good 50 percent of the value (of formal inservice education activities) and value enough.

Consequently, they like activities to be arranged in a manner that permits and encourages social contacts.

I see inservice activities as the perfect opportunity to talk with other teachers in the same field as you are. I see that as a significant benefit.

Too, attending activities reminds them that they belong to a profession.

Despite the fact that often the offerings are not as useful as one would hope, I think there is an aura of fraternity, of fellowship, that surfaces at this type of gathering which makes it worthwhile, regardless of what happens . . . you go about your working day and you forget that you belong to a profession.

F. *To obtain recognition that a worthwhile job is being performed.* Teachers suffer from a lack of positive feedback concerning the importance and effectiveness of their performance. Consequently, they seek presentations which give them a



mental pat on the back, supply the desired recognition and encourage them to return to the classroom and persevere with their teaching assignment.

It's confidence building in some ways. It's hard to get recognition (in teaching), it's one of the problems of the profession. The kids don't say it, adults don't see you perform and you just don't know . . . inservice serves that type of purpose, it's confidence building.

To a great extent we are appreciated by the community but are not told about it, we just hear the negative and it's nice to be told. Keynote speakers tell us that . . . and we take the pats on the back and say maybe we're not doing such a bad job.

G. *To enjoy a mentally stimulating break from routine.* It is difficult for teachers to escape mentally from the confines of the classroom, and inspiring, thought-provoking presentations which permit teachers to extend their mental horizons beyond the four walls of their own classroom and subject specialization are welcomed. Such presentations help them to keep their own classroom and its concerns in perspective and they return rejuvenated and encouraged.

Inservice on any topic is refreshing, a change of pace, a little spark.

Going to meetings, workshops can open up new vistas.

From a shot-in-the-arm session you get a sparkle and the kids benefit.

### *Group Three: Formal Inservice Education Activities of a Voluntary Nature*

These activities usually are organized by the school, the school system, the specialist councils of the teachers' professional association or the university, and may be held during school hours, in the evening, on the weekend or during school holidays. For this group of activities, teachers do not believe that benefits will accrue regardless of topic or presenter, as they do for mandatory formal activities that do not accompany a compulsory program change. Although teachers may attend these activities for any of the reasons listed in Figure 1, in general they seek information and materials which will help them with their present teaching assignment. They prefer to have these activities scheduled within normal school hours since they do not feel they have surplus time or energy for activities held out of working hours.

The criterion they use to determine if they should attend these activities is "will the rewards equal the sacrifices?" If the activity is held during the school day, the sacrifices or costs incurred by attending are the loss of teaching time and the work of preparing for a substitute teacher. If the activity attended does not provide them with the anticipated rewards, they suffer feelings of guilt because they consider that they have unnecessarily deprived their students of valuable teaching time.

You should leave inservice feeling excited and wanting to go back to the classroom and try it out yet the last inservice here I left and felt badly about the waste of time. I would rather have been teaching.

If the activity is held outside of normal school hours, attending means forgoing part of their leisure time and they want reassurance that the benefits derived will compensate this loss.

Depends on the value of the session. If I can be guaranteed it will be worthwhile, then I don't mind evening or weekend.

That teachers believe they do not have the time or energy to attend activities held out of school hours agrees with McClusky's (1973) concept of margin. McClusky (1973) considers that a necessary condition for learning is the availability of margin, or surplus power or energy left over after a person has handled his load. Given society's never-ending expectations of the educational system, the knowledge explo-



sion and the increasing number of teachers who are in the middle years, it is not surprising that many teachers believe that by the end of the school day, week or year, they do not have any surplus energy for formal inservice education activities.

I don't mind devoting some of my own time but when you ask me to go at four o'clock, I don't have much to put into it . . . teachers are worn out by the end of the school day.

As one teacher explained:

Teachers fear they (inservice activities) are a waste of time because they frequently have been. They want a guarantee that when they spend time it will pay off . . . also there's the problem of conservation of resources, a personal thing: How much can I spend and still survive as a human being? They do have the time and energy if they can be guaranteed a payoff—no time means no time to waste on inservice if they are guaranteed only a 20 percent chance of it being worthwhile.

#### *Group Four: Non-Formal Inservice Education Activities*

Non-formal inservice education activities are ongoing, often casual, involvements which can occur at any time and in any location. Usually they are teacher initiated and involve an individual teacher or a small group of colleagues. The most frequently cited non-formal activity is daily in-school contacts with colleagues, in particular with those colleagues who have the same subject specialization.

Yet it goes on all the time—when you think about it, professional development goes on all the time in the staffroom.

Peer-based consultation, one-to-one or small group, is more important than formal inservice workshops.

I pick up more in the staffroom during the course of the year than at workshops.

Other non-formal activities are reading professional literature, involvement with new programs, having student teachers and watching others teach, and participating in extra-curricular activities that have spin-offs to the classroom.

Reading books and periodicals is of great value. I've become very familiar with the ATA library . . . I try to visit there once every three months.

My professional development comes from working with colleagues in school and from my choice of extra-curricular activities. Reading too . . . a lot comes through the introduction of new programs.

You can get professional growth by watching other teachers in action in their schools. Also, working with student teachers is a professional activity and every teacher could benefit from it.

Teachers indicated that they regularly use non-formal inservice education activities to keep current, obtain additional classroom-related information and materials, and find solutions to classroom-related problems. Thus it appears that, in general, it is from their involvement in these activities rather than formal inservice education activities, that teachers obtain information and materials which they will use to effect major changes in their classroom behavior and so improve their effectiveness. Since participation in these activities directly affects their performance in the classroom, teachers consider these involvements to be, at a minimum, worthwhile if not essential. The importance to teachers of non-formal activities, in particular in-school contacts with colleagues who have the same areas of specialization, are revealed by the comments of teachers who are the only specialists in their subject area in the school.

What I do enjoy (at inservice activities) is meeting with other librarians in the district because the business of being a librarian is relatively lonely. There may be five social studies teachers in the school so that if you have a problem you can kick it around a bit, but when

you are a one-teacher specialist such as the librarian or the home economics teacher, you are alone.

### *Conclusions, Recommendations and Implications*

#### *Conclusions*

Traditionally, it has been assumed that by participating in formal inservice education activities teachers will obtain information which they will use to bring about an immediate change in their classroom behavior and so maintain or improve their effectiveness. According to the findings of this study, such an assumption is not always accurate. Teachers do endeavor to keep current and provide their students with the best learning experiences possible, but they use a variety of methods to attain this goal, and consider in-school contacts with colleagues, rather than involvement in formal inservice education activities, to play a vital role.

Thus, based on the findings of this study, it appears that formal inservice education activities which do not accompany a mandatory program change are not an appropriate method to use to effect a major change in the classroom behavior of teachers. These activities, however, do serve a useful purpose, as attending them permits teachers to satisfy certain basic and essential needs in the affective and lower level cognitive domains. Consequently, such activities should be considered worthwhile and should continue to be organized for teachers.

Formal inservice education activities can, however, make a more important contribution than simply permitting teachers to satisfy certain basic needs. But first, there must be a long-term plan for teacher professional development. This should take into consideration the principles of adult learning, the developmental stages of adults, the nature and pattern of teachers' professional needs, and should include all of the parties involved in teacher professional development—the individual, the employer, the professional association and the university. Formal inservice education activities would comprise but one component of the plan, and would be organized in a manner which acknowledged and provided for individual as well as subject area differences. It should ensure that teachers have choice, choice among activities and also choice as to whether or not to attend. It should involve teachers, not relegate them to passive roles, should recognize that they bring to sessions valuable knowledge and experience, and should permit them to enjoy and benefit from social contacts with colleagues.

It is interesting to note that school systems traditionally have put their money, time and energy into organizing formal inservice education activities and have often given little encouragement to, or ignored, teachers' involvement in non-formal activities. Yet given the findings of this study, it appears that it is through their involvement with non-formal, not formal, activities that teachers are obtaining knowledge which they will use to bring about changes in their classroom behavior; they attend formal activities for other, but to them very important, reasons. If teachers are to be the beneficiaries of inservice education, it could be argued that it is their perspective which should determine the nature of inservice education.

#### *Recommendations and Implications*

*Recommendations for School Systems.* The district or central office of the school system should continue to organize formal inservice education activities for teachers and also provide some funds for individual professional development.



Teachers should be given release time to attend activities held during school time.

- The district office of the school system should acquaint the public with the nature of, and need for, continuing teacher professional development. In particular, the public should be made aware of the fact that teachers are involved in ongoing continuing education experiences and that many of the activities which contribute to an improvement in teacher effectiveness are of a non-formal, often apparently social, nature. Too, that on days when students are released from classes, teachers can be engaged in meaningful learning experiences without attending group activities held in local schools.
- All formal inservice education activities should be held within the school day, week and year. The school system should organize such activities or should permit teachers to attend such activities organized by others. Part, if not all, of the cost should be borne by the school system.
- When introducing compulsory program changes, the school system and/or the department of education should ensure that such program changes are accompanied by formal inservice education activities held during the school day, preferably with students released from classes. The presentations should be given by classroom teachers with expertise in the area who can help teachers interpret and implement the new curriculum and who can suggest appropriate classroom materials.
- The formal inservice education activities organized for teachers should include the following types of presentations:
  - a. Practical, classroom-related types of presentations preferably given by those with recent classroom experience in the area. The ideas presented should be capable of being immediately adopted by the teacher on his return to the classroom and should require only those materials that are presently available in the schools. The preferred format for such activities is small group, hands-on activities which permit teachers to become involved and share ideas with other participants. Handouts are appreciated as they remind teachers of the information presented.
  - b. Information sessions that acquaint teachers with the latest developments in the field of education.
  - c. Positive, stimulating, thought-provoking presentations that enhance teachers' self-concept. Such presentations may relate directly to the field of education or may be of a more general nature.
- Since in any group of teachers there will be those who are novices in an area and those with a great deal of experience, subject matter presentations should be offered at a variety of levels to cater to the different needs of teachers at different stages in their careers.
- Since teachers have indicated their preference for presenters who are classroom teachers and pointed out the lack of qualified presenters, the school system should encourage and reward, in some way, those teachers who express an interest in giving occasional presentations and who are considered to have the necessary expertise. It is suggested that these rewards need not necessarily be of a financial nature but could take the form of release time, recognition, reimbursement for expenditures, and so forth.
- It is suggested that the school system should not rely too heavily on imported speakers who have no knowledge of the local scene but simply give a prepared presentation and leave.
- When organizing activities, the school system should take the time of year into consideration. Teachers have indicated that at certain times of the year they prefer more subject-oriented presentations whereas at other times of the year they desire thought-provoking, mentally stimulating presentations.
- Formal inservice education activities should be organized in a manner which deliberately encourages interactions of a social nature.
- The formal inservice education activities organized by the school system should follow and support a long-term plan for teacher professional growth. The weakness of a "one of" workshop is not so much the fact that it represents only one presentation on a topic, but rather



that the single presentation has nothing whatsoever to do with the previous or the following session.

- The importance of non-formal inservice education activities should be acknowledged by the school system and involvement in these types of activities should be encouraged.
- Teachers should have the freedom to choose if, how, and when to participate in formal and non-formal inservice education activities. In particular, on school days when students are released from classes, they should have the freedom to participate in whatever types of activities they consider will best benefit both themselves as professionals and their students.
- If the school system wishes to have teachers adopt an innovation, the following suggestions, based on the comments of the teachers included in the study, should be taken under consideration:
  - a. Teachers must be exposed to the innovation on several occasions over an extended period of time in order to permit them to become thoroughly acquainted with, and accept, the proposed innovation.
  - b. The school system must not only expose teachers to the innovation but also must take steps to encourage teachers and affected others to adopt the innovation. Teachers included in the study indicated that it is much easier to adopt an innovation if other members of the staff also are willing to change. The present common practice of exposing only one member of the school staff to an innovation and expecting that individual to persuade the rest of the staff to change, is unrealistic. As one teacher said, "when others on staff are unfamiliar with a new idea, it fizzles out."
  - c. The school system must realize that even when teachers wish to change their classroom behavior, usually they do not do so until "next time around," thus the timing of the exposure to the innovation is important.
- Since teachers have indicated that they have no time and little interest in being involved in planning formal inservice education activities, the present practice of offering concurrent sessions is an appropriate way of attempting to meet the needs of different groups of teachers. When organizing such activities, it should be remembered that teachers desire not only different topics but also different levels of subject matter presentations, and that teachers wish to have the choice of participating in formal or non-formal activities.
- Many of the teachers included in the study used terms which suggest that they perceive themselves as being isolated and powerless members of a large, impersonal organization. It is suggested that the school system should make every effort to ensure that teachers are, and perceive themselves to be, important, contributing members of the organization who have some control over the various aspects of their professional life.
- Since many teachers mentioned feelings of isolation and indicated that they considered in-school contacts with colleagues to be an important form of non-formal inservice education, it is suggested that the school system should deliberately foster such teacher-initiated consultations.
- A special effort should be made by the school system to cater to the needs of teachers who are the only subject specialists in their schools, as such teachers feel particularly isolated and neglected. They are deprived of ongoing, in-school contacts with colleagues in the same area of specialization and often their needs are overlooked at formal activities.
- School systems should consider introducing voluntary teacher-to-teacher networks similar to those described in *IMPACT II* (Mann, 1982). The approach to inservice education described in *IMPACT II* capitalizes on an important finding of this study, namely, that in-school consultation with colleagues contributes to the professional growth of teachers and the alleviation of their concerns.
- It is suggested that the school system could use the instrument developed by the researcher (Figure 1) to ascertain the needs that teachers desire to have satisfied by attending the various groups of inservice activities, and the value attributed to each group. Such information could help avoid a mismatch between the goals of organizers and participants.

- At present, formal inservice education activities do not appear to be evaluated, at least not in any strict sense of the term. If, in the future, an evaluation of these activities is considered to be desirable, it is suggested that since many of the benefits derived by participants are of an intangible nature, both traditional and non-traditional evaluation methods should be used.

*Recommendations for Universities.* Some teachers considered the university to play a vital role in their professional development; but many considered it to be a degree-granting institution, a “salary upgrading place,” which did not play a significant role. It is a place “for me to use in any way I want,” to which they turn when they want in-depth knowledge or a guest speaker. But, when they seek information which will help them improve their performance in the classroom, they turn elsewhere because, in their eyes, professors lack credibility—they are too far and too long removed from the world of the school and the classroom teacher.

Don't mention the university in regards to inservice as it's the ivory tower; it's years since they've been in the classroom. I'd go to university to be brought up-to-date on things in fantasyland.

The presentations are fairly useless if you go back (to school) with the intent of using these ideas on Monday morning.

The most frequently made recommendation was that professors in the faculty of education should have recent and regular exposure to the world of the classroom teacher.

My big gripe . . . is that many people in the faculty of education have not been in the classroom for 20 years . . . it's not the same just coming to observe a student teacher. So if you train teachers, after 5 years you should be back in the classroom for one day a week. You have to teach these people—be responsible for a class—to see what it's like now. Professors have a total misconception of what the classroom is like now.

Also, it was suggested that the university could play a more positive and powerful role in developing the preservice-inservice continuum.

At university (preservice), no one talked to us about inservice, we just saw notices posted. I didn't know if you were supposed to go or not. I went to a couple of Saturday morning sessions and found them quite valuable but wished my friends had been there as I felt quite inadequate with teachers of eight to ten years' experience. The university should stress attending inservice activities.

Following are recommendations for the university, based on the findings of this study which, it is suggested, would permit the university to play a more active role in the continuing education of teachers.

- The university should offer a wide variety of short-term credit and non-credit, off-campus and on-campus courses similar to those evaluated by Unruh (1981) which are organized in accordance with teachers' perspective on inservice education.
- The university should take steps to deliberately foster the development of a preservice-inservice continuum. In particular, preservice education students should be exposed to the concept of lifelong continuing teacher education and the various types of formal and non-formal ways in which teachers can continue their inservice education. Undergraduate education students should be encouraged to attend formal inservice education activities as part of their teacher education training; attending such activities should help to develop a feeling of professionalism and collegiality.
- As Kersh (1979, p. 45) states, “college faculty still act as if inservice is ‘doing something to somebody or running something for somebody.’” The university should encourage an ongoing, two-way exchange of ideas between university faculty and teachers in the classroom.
- University professors in the faculty of education should be involved in a continuing professional education program, a major component of which would be regular, in-depth classroom experiences.



*Towards the Generation of Substantive Theory*

Glaser and Strauss (1965, 1967) advocate the discovery of theory from data systematically obtained from social research, which they call grounded theory, as opposed to theory generated from a priori assumptions. They consider that grounded substantive theory:

can give participants in a situation a broader guide to what they already tend to do, and perhaps help them to be more effective in doing it. (Glaser & Strauss, 1967, p. 248)

Based on the findings, following is an attempt to predict and explain teachers' attitudes and behavior in relation to inservice education.

- Teachers obtain information which they use to effect major changes in their classroom behavior and so improve their effectiveness as teachers from their involvement in a variety of non-formal inservice education activities, not from their involvement in formal inservice education activities, apart from those activities which accompany a compulsory program change.
- Most teachers are not dissatisfied with their present classroom practices and attend formal inservice education activities which do not accompany a compulsory program change in order to obtain confirmation that their present classroom practices are acceptable and appropriate, not to obtain information which they will use to effect a major change in their present classroom behavior.
- Teachers consider formal inservice education activities that do not accompany a compulsory program change to be information sessions not sessions designed to immediately change their classroom behavior.
- Teachers consider formal inservice education activities that do not accompany a compulsory program change to have little relationship to their day-by-day classroom activities. They attend formal inservice education activities that do not accompany a compulsory program change in order to satisfy lower level cognitive needs or needs in the affective domain.
- Teachers readily adopt new techniques and materials which require no major changes in their present classroom behavior or ideology. They adopt innovations which require a major change in their classroom behavior in crisis situations only, or when they are compelled to do so because of a compulsory program change.
- Teachers do not voluntarily adopt an innovation which requires a major change in their classroom behavior until they have been exposed to the innovation on several occasions over an extended period of time.
- Teachers use a cost-benefit formula to determine whether or not to attend voluntary formal inservice education activities.
- There is an identifiable pattern to the inservice education needs of teachers which changes over time as the individual moves along his career path from neophyte to experienced teacher. When a teacher changes grades, area of specialization or is promoted, however, initially in the new position he tends to function as a neophyte.
- Teachers who have satisfied most of their inservice education needs at all levels from deficiency to growth seek new challenges by changing grades, areas of specialization or seeking promoted positions.
- Teachers who are deprived of ongoing, in-school contacts with colleagues who have the same area of specialization suffer from extreme feelings of isolation and are deprived of an important source of information for improving their classroom effectiveness.
- One cannot truly understand the world of another unless one lives in that world. Consequently, in order for university professors to be able to offer classroom-related assistance to teachers, they must regularly live in the world of the classroom teacher.



## Appendix

### *Final Categories Developed From the Data*

The following four elements of a perspective, as identified by Becker, Geer, Hughes, and Strauss (1961, p. 436), form the conceptual framework used to portray teachers' perspectives on inservice education:

- a. a definition of the situation in which the actors are involved,
- b. a statement of the goals they are trying to achieve,
- c. a set of ideas specifying what kinds of activities are expedient and proper,
- d. and a set of activities or practices congruent with them.

The element is identified to the left of the category number:

- (a) 1. Teachers' definitions.
- (b) 2. Attitude of teachers towards their own professional development.
- (c,d) 3. Attitudes of teachers towards formal inservice education activities.
- (c,d) 4. Teachers' reasons for attending formal inservice education activities that are not a part of a compulsory program change.
- (c,d) 5. Teachers' views on topics offered at formal inservice education activities.
- (c,d) 6. Changes in attitude over time towards formal inservice education activities.
- (c,d) 7. Teachers' views on attendance and timing of formal inservice education activities.
- (c,d) 8. Teachers' views on presenters at formal inservice education activities.
- (c,d) 9. Teachers' criticisms of formal inservice education activities.
- (c,d) 10. Non-formal forms of inservice education.
- (c,d) 11. Teachers' comments relating to inservice education and the role of district office.
- (c,d) 12. Teachers' comments relating to inservice education and the role of the university.

## References

- Alberta Teachers Association. (1981). Committee proposes solutions to inservice problems. *ATA News*, 15(10), 1.
- Bacon, R. (1980). Inservice in Britain: INSET or inservice education and training? *Elements*, 12(3).
- Becker, H. S., Geer, B., Hughes, E., & Strauss, A. (1961). *Boys in white*. Chicago: University of Chicago Press.
- Blumer, H. (1969). *Symbolic interactionism: Perspective and method*. Englewood Cliffs, NJ: Prentice-Hall.
- Blumer, H. (1978). Society as symbolic interaction. In A. Wells (Ed.), *Contemporary sociological theories*. Santa Monica, CA: Goodyear.
- Bogdan, R. C., & Biklen, S. K. (1982). *Qualitative research in education: An introduction to theory and methods*. Boston: Allyn & Bacon.
- Brayne, R. (1980). The principal and teacher professional development—some strategies. *Studies in Educational Administration*, 21.
- Bruce, M. G. (1980). Inservice education grows massively in Western Europe. *Phi Delta Kappan*, 62(2), 140-141.
- Bruyn, S. T. (1966). *The human perspective in sociology*. Englewood Cliffs, NJ: Prentice-Hall.
- Cicourel, A. V., & Kitsuse, J. I. (1963). *The educational decision-makers*. Indianapolis: Bobbs-Merrill.
- Clifford, G. J. (1973). A history of the impact of research on teaching. In R. M. W. Travers (Ed.), *Second handbook of research on teaching*. Chicago: Rand McNally.

- Cooper, J. O., & Hunt, K. P. (1978). A cooperative approach to inservice training. *Viewpoints in Teaching and Learning*, 54(4), 61-69.
- Cruikshank, D., Lorish, C., & Thompson, L. (1979). What we think we know about in-service education. *Journal of Teacher Education*, 30(1), 27-31.
- Denzin, N. K. (1970). *The research act*. Chicago: Aldine.
- Dexter, L. A. (1970). *Elite and specialized interviewing*. Evanston: Northwestern University Press.
- Douglas, J. D. (1976). *Investigative social research*. Beverly Hills: Sage.
- Fennell, B. H., Hill, W. L., & Thiessen, S. J. (1980). *Teacher inservice training costs: A staff study*. Alberta: Alberta Education.
- Frandsen, P. E. (1980). Continuing education for the professions. In E. J. Boone, R. W. Shearon, & E. E. White (Eds.), *Serving personal and community needs through education*. San Francisco: Jossey-Bass.
- Fullan, M., & Pomfret, A. (1977). Research on curriculum and instruction implementation. *Review of Educational Research*, 47(1), 335-397.
- Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press.
- Glaser, B. G., & Strauss, A. L. (1965). Discovery of substantive theory: A basic strategy underlying qualitative research. *American Behavioral Scientist*, 8, 5-12.
- Glaser, B. G., & Strauss, A. L. (1966). The purpose and credibility of qualitative research. *Nursing Research*, 15(1), 56-61.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago: Aldine.
- Goodman, P. S. (1977). *New perspectives on organizational effectiveness*. San Francisco: Jossey-Bass.
- Guba, E. G. (1981). *Criteria for assessing the trustworthiness of naturalistic inquiries*. Paper presented at the University Council for Educational Administration Graduate Student Seminar, University of Alberta, Edmonton, May 3-6, 1981.
- Haller, E. J. (1979). Questionnaires and the dissertation in educational administration. *Educational Administration Quarterly*, 15(1), 47-66.
- Harrington, F. H. (1979). *The future of adult education*, San Francisco: Jossey-Bass.
- Houle, C. O. (1963). *The inquiring mind*. Madison: University of Wisconsin Press.
- Houle, C. O. (1980). *Continuing learning in the professions*. San Francisco: Jossey-Bass.
- Howey, K., & Joyce, B. (1978). A data base for future directions in inservice education. *Theory Into Practice*, 17(3), 206-211.
- Kersh, B. Y. (1979). Needed: New ways of thinking and acting in staff development. *Journal of Teacher Education*, 30(1), 44-46.
- King, J. C., Hayes, P. C., & Newman, I. (1977). Some requirements for successful inservice education. *Phi Delta Kappan*, 58(9), 686-687.
- Mann, D. (1982). The impact of IMPACT II. *Phi Delta Kappan*, 63(9), 612-614.
- McClusky, H. (1973). An approach to a differential psychology of the adult potential. In M. Knowles, *The adult learner: A neglected species*. Houston: Gulf.
- Richey, H. G. (1957). Growth of the modern conception of in-service education. In N. B. Henry (Ed.), *In-service education for teachers, supervisors, and administrators*. Chicago: University of Chicago Press.
- Spradley, J. P., & McCurdy, D. W. (1972). *The cultural experience: Ethnography in a complex society*. Chicago: Science Research Associates.
- Stephens, T. M., & Hartman, A. C. (1978). Inservice education: Its murky past and uncertain future. *Viewpoints in Teaching and Learning*, 54(4), 1-3.

- Turner, B. A. (1981). Some practical aspects of qualitative data analysis: One way of organizing the cognitive processes associated with the generation of grounded theory. *Quality and Quantity*, 15, 225-247.
- Unruh, W. R. (1981). *An evaluation of the design, development and implementation of several course formats for teacher in-service in education*. Calgary: Faculty of Education, University of Calgary.
- Webb, S., & Webb, B. (1968). *Methods of social study*. New York: Kelley. (Original work published in 1932).
- West, W. G. (1975). Participant observation research on the social construction of everyday classroom order. *Interchange*, 6(4), 35-43.
- West, W. G. (1977). Participant observation in Canadian classrooms: The need, rationale, technique, and development implications. *Canadian Journal of Education*, 2(3), 55-74.
- Wood, F. H., & Thompson, S. R. (1980). Guidelines for better staff development. *Educational Leadership*, 37(5), 374-378.



LEE GUNDERSON

*The University of British Columbia*

## One Last Word List

*This article begins with a discussion of large corpus and pedagogical word lists. Pedagogical lists, generally developed from large corpus lists, are constructed to represent significant portions of the vocabulary in school texts. Teachers often teach the words on pedagogical lists so their students will know on sight most of the words they will encounter in written discourse. Over the last two decades many lists have been published as improvements to the "Dolch list of 220 words," the most widely used list, each claiming they represent more of the vocabulary in selected texts.*

*Two pedagogical lists are developed here, one based on the words produced by 80 second, fifth, eighth, and eleventh grade subjects writing three different compositions and the other produced by only the second grade subjects. The first list is as successful as well known published lists, including the Dolch, in accounting for substantial portions of the vocabulary of beginning reading texts. The second grade list, developed from a corpus of only 2,000 words, represents over half of the vocabulary in beginning reading texts.*

*The article concludes with the suggestion that researchers have done enough as far as pedagogical word lists are concerned; they are all equally successful in representing the vocabulary of beginning reading texts. The task remaining is to investigate the vocabulary of individual beginning reading texts, vocabularies that are significantly different than those represented by either large corpus or pedagogical word lists.*

### Introduction

Over the last decade there has been active interest in two types of word lists, large corpus word lists and pedagogical word lists. The first is based on a rank ordering of the words in a sample while the second is a limited pedagogical list constructed to represent significant portions of textbook vocabulary. The large corpus word lists are of interest to researchers attempting to understand more about our language. Teachers teach the words on pedagogical lists because they feel they are important for their students' success in reading. One of the most influential pedagogical lists is the Dolch (1936).

---

Lee Gunderson (PhD, University of California, Berkeley) is an Assistant Professor in the Department of Language Education at the University of British Columbia. His interests are in the area of first and second language reading acquisition.

Dolch produced "A Basic Sight Vocabulary of 220 Words, Comprising All Words, Except Nouns, Common To The Word List of the International Kindergarten Union, The Gates List, And The Wheeler-Howell List." Dolch added some words that were not common to all three lists because:

Many of these words obviously belong with others on the basic list. "Which" belongs with "who" and "that," "done," and "goes" belong with "did" and "go," "start" belongs with "stop," and "write" with "read." (Dolch, 1936, pp. 457, 459)

In 1960 Dolch took several 1000-word samples from reading, math, geography and history texts from grades one to six and found that the words on his list represented a great deal of the vocabulary encountered; for instance, 70 percent of the vocabulary in grade one reading texts to 52 percent of the vocabulary in grade six history texts.

Dolch suggested his list be taught in order to help students develop a sight vocabulary useful in reading all subject areas. He stated, "It is not claimed that the basic list of 220 words includes all the words that the elementary school pupil should know by sight, the claim is only that he should at least know these" (Dolch, 1936, p. 459).

As Hillerich (1974) pointed out, "most educators accept the need for a basic list" (p. 354). For years teachers have used the Dolch as a basic list of sight words to be taught. Occasionally primary teachers teach the Dolch words to their students *before* they involve them in formal reading programs. Some use it as a diagnostic and/or criterion measure since they consider the ability to recognize the 220 words by sight to be a significant indicator of reading achievement. Yet there are critics who maintain that the Dolch is outdated.

Johnson (1971a) stated, "The Dolch List, as a corpus, has outlived its usefulness" (p. 449). Johnson compared the Dolch list with the Kucera and Francis (1967) list which was developed from 500 samples of 15 categories of adult literature. He found that 37 percent (82 words) of the Dolch list were not in the top 220 words on the Kucera-Francis corpus. He concluded that the Dolch list should be replaced by the highest frequency Kucera-Francis words. He responded to the criticism that the Kucera-Francis is based on an adult corpus rather than a children's corpus by stating, "*is it not more important for primary reading materials to reflect what exists in present-day American English, than for sight words lists to reflect what occurs in beginning reading material?*" (p. 457).

Johnson (1971a) constructed his own word list on the basis that items would be: 1) in the top 500 on the Kucera-Francis list and 2) used at least 50 times by kindergarten and first grade students in a study of spontaneous oral language (Murphy, 1957). His list contained 306 words. Hillerich (1974) believed this list "tended to overkill with adult words" (p. 359). Hillerich's own list (1974) consisted of 240 starter words that represented a basic language arts vocabulary. He found that 62 of the Dolch words were not found on his list. He noted that the Dolch may be inappropriate rather than outdated.

Johns (1976a), in a series of studies, developed the Revised Dolch List (RDL). First, he eliminated the 31 words not common to at least three of four lists: 1) the Carroll, Davies, and Richman (1971), (a study of 5,088,721 words occurring in grades three to nine textbooks), 2) the Durr (1973) (the 188 most frequent words in primary library books), 3) the 500 most frequent Kucera and Francis (1967), and 4) the 727 words occurring at a minimum frequency of 50 in the Murphy (1957) study (Johns, 1974). Johns (1976b) added words common to at least three of the



lists mentioned above and produced the RDL of 226 words. Johns (1976a) reported that his list represented significant amounts of the vocabulary of basic readers in four reading series. He also noted that his list represented significantly more of the vocabulary in text than did the Dolch.

The Dolch has its supporters, however. Lowe and Follman (1974) compared the first 150 Dolch words with four other lists and found a high degree of similarity. They suggested the first 150 words can be used "without reservation or limitation" (p. 44). Mangieri and Kahn (1977) analyzed the vocabulary in four basal reading series by taking 3,000 word samples and comparing them to the Dolch. They found the Dolch accounted for 58 to 76 percent of the vocabulary in initial reading texts. Seventy percent of the Dolch words occurred at least once in the basals they studied. Mangieri (1978) showed that significant portions of the vocabulary in four basal reading series were accounted for by the Dolch list (66 to 72 percent).

Thomas (1972) produced a list based on the writing of students in grades one to six in Alberta. He randomly selected about two hundred written products from each grade level resulting in a total of 1,287 compositions and a cumulative 117,878 running words. Thomas' rationale for developing a Canadian list was that the needs of educators, textbook authors, and students would be better served by such a list than by lists developed by American researchers using American subjects. He states, "textbook writers have had to rely heavily upon American vocabulary lists which may not necessarily be meeting the particular regional needs in Canada" (Thomas, 1972, p. 243). He infers that the vocabulary of students in Canada is different than the vocabulary of American students. He concluded, however, that a small number of high frequency words constitutes the majority of the vocabulary in students' writing; differences in vocabulary were apparent in relatively low frequency items.

The purpose of this study was to develop a word list from a small sample of words from a unique population of subjects. It was hypothesized that such a list would be similar to other published lists. If the hypothesis was substantiated, the generated list would account for as much of the vocabulary in initial reading texts as other pedagogical lists. Further, findings would show the similarities between the vocabulary of the subjects and the vocabulary of initial reading texts. If confirmed, the hypothesis would suggest that regional vocabulary needs are not met by simply producing regional vocabulary studies in the traditional manner. Such a finding would suggest that alternate methods of studying vocabulary will have to be developed in order to meet the regional needs in Canada.

### *Method*

#### *Subjects*

One objective of the study was to collect vocabulary samples from a population that was obviously unique. With this goal in mind, subjects from a large metropolitan school district in the Western United States were selected randomly from three high schools, four middle schools, and ten elementary schools at grades two, five, eight, and eleven. Each of the participating schools had mean scores in total reading on the Comprehensive Test of Basic Skills (C.T.B.S.) below the 50th percentile. A total of 311 eleventh grade subjects, 321 eighth grade subjects, 309 fifth grade subjects and 268 second grade subjects wrote one or more of the three tasks. A random sample of 20 subjects at each grade level was selected from the 43 percent who had written all three products. The sample consisted of approximately 17 percent



Spanish surname, 24 percent Black, 12 percent Filipino, 20 percent Asian, 18 percent White and 9 percent other (e.g., East Indian). Approximately 22 percent were English as a Second Language (E.S.L.) students. These percentages are close to the population percentages of the schools in the sample and the district at large.

### Procedure

Each subject wrote three different essays: a description, a narration, and an argument (writing prompts are shown in the appendix). The assignment was printed on the first page of the test booklet and the teacher read it aloud. Subjects were encouraged to try to spell words correctly, but were given no help. A total of forty minutes was given for each writing task.

Each of the 240 compositions was recorded in permanent computer files. Proper nouns and abbreviations were included as written. Illegible items were ignored. Spelling errors were listed separately, but not included in the recordings. Two error types were excluded; 1) obvious errors (i.e., wuz) and 2) context errors (i.e., "I no you were here."). A program entitled, "A Word Counting and Frequency Analysis Program" (Miller, 1975) was used to count words. In this case a word was defined as any string of characters bounded by blank spaces.

### Results

Analysis of the second grade data showed there were 2,011 tokens (total words) and 361 types (different words) for a type/token (T/T) ratio of .18. The fifth grade sample contained 3,178 tokens and 582 types (T/T=.18). The eighth and eleventh grade subjects produced 4,145 and 726 (T/T=.18) and 6,951 and 1,161 (T/T=.17) respectively. When combined, the corpus produced 16,285 tokens and 1,675 types (T/T=.10).

The top 250 words were selected and a list compiled. Since the absolute frequency of words 251 to 253 was the same as the frequency of 250, they were included as well. Inspection of the 253 words showed there were some specifically topic prompted items, i.e., the word *special* in the description task. Eleven words were eliminated as being specifically topic-prompted. The 242 words can be seen in Table 1.

The list contains many inflected forms, i.e., *walk*, *walks*, *walked*, *walking*. Unlike the Dolch list, it also contains nouns such as *mother*, *father*, *birthday*, *teacher*, etc.

### Pedagogical List Results

The Inner-City list was compared with four pedagogical lists: Dolch (1936), Johns RDL (1976), Hillerich (1974), and the Durr (1973). The hypothesis that the I.C. list would be similar to other pedagogical lists was confirmed. The I.C. and Dolch lists share 144 words (60 percent of the I.C.), Johns RDL 155 (64 percent), Hillerich 176 (72 percent) and the Durr 146 (60 percent). The I.C. list accounts for 65 percent of the Dolch, 69 percent of the Johns, 73 percent of the Hillerich and 78 percent of the Durr.

The 100 most frequent words on the Thomas (1974) Canadian word list representing 58 percent of his total word count and the I.C. list shared 94 words. The six Thomas words not found on the I.C. list were *dog*, *people*, *two*, *time*, *dad*, and *once*.

TABLE 1  
RANK ORDER LIST OF INNER-CITY WORDS

|          |         |           |            |          |          |
|----------|---------|-----------|------------|----------|----------|
| the*     | then*   | other     | talk       | say*     | again*   |
| to*      | get*    | has       | their*     | wanted   | an*      |
| and*     | what*   | into      | around*    | way      | call*    |
| you*     | there*  | nice      | ever       | back     | end      |
| he*      | about*  | too*      | everything | can't    | happy    |
| is*      | all*    | play*     | inside     | doesn't  | helps    |
| I*       | good*   | hope*     | looked     | family   | knew     |
| a*       | this*   | much*     | most       | keep*    | little*  |
| me*      | see*    | some*     | over*      | kind*    | next     |
| it*      | bad*    | take*     | since      | lot      | pick*    |
| in*      | out*    | father    | try*       | name     | problem  |
| my*      | will*   | something | walking    | only*    | same     |
| she*     | boy     | us*       | come*      | them     | seen     |
| was*     | just*   | got*      | didn't     | together | still    |
| for*     | saw*    | could*    | found*     | did*     | sure     |
| of*      | up*     | how*      | house      | doing    | walk*    |
| that*    | they*   | even      | well*      | eat*     | won't    |
| your*    | or*     | help*     | work*      | goes*    | year     |
| his*     | had*    | no*       | after*     | last     | big*     |
| so*      | want*   | now*      | each       | need     | class    |
| with*    | go*     | life      | give*      | our*     | close    |
| have*    | please* | might     | problems   | she's    | coming   |
| when*    | think*  | care      | said*      | should   | door     |
| very*    | always* | he's      | told       | two      | down*    |
| but*     | school  | it's      | also       | use*     | funny*   |
| can*     | who*    | more      | birthday   | walked   | gets     |
| her*     | from*   | by*       | first*     | which*   | gives    |
| we*      | love    | fun       | look*      | another  | left     |
| like*    | fine    | home      | sick       | before*  | let*     |
| on*      | tell*   | make*     | teacher    | brother  | living   |
| because* | would*  | sometimes | that's     | food     | made*    |
| not*     | day*    | came*     | years      | its*     | makes    |
| do*      | at*     | does*     | am*        | looks    | night    |
| if*      | why*    | were*     | any*       | lots     | somebody |
| don't*   | man     | I'm*      | anything   | many*    | wish     |
| him*     | as*     | been*     | away*      | may*     | write    |
| are*     | things  | money     | long*      | off*     | wrong    |
| know*    | went*   | open      | looking    | sister   |          |
| be*      | took    | put*      | mom        | than     |          |
| mother   | going*  | really    | never*     | thing    |          |
| one*     | opened  | right*    | old*       | yourself |          |

\* Included on Dolch list

Large Corpus Results

The hypothesis that the I.C. list would be similar to other published lists was substantiated. Correlation coefficients were calculated for the association between I.C. frequencies and frequencies on the Durr (1973), the Carroll, Davies, and

Richman (1971) and the Rinsland (1947), a study of the 6,012,359 words written by students in grades one to eight. The correlations are Durr .83, Rinsland .88, and Carroll, Davies, and Richman .73.

### *Initial Reading Text Results*

The hypothesis that the I.C. list would be as successful as the others in accounting for the vocabulary in initial reading texts was confirmed. The Macmillan (1980) "series r" preprimers and primers (levels 4, 5, 6 and 7) were analyzed. It was found they contained 4,620 words in connected discourse with 323 types for a T/T of .07. Considerable portions of the vocabulary were accounted for by the different lists; 129 I.C. words accounted for 69 percent, 115 Dolch words accounted for 69 percent, 115 Durr words accounted for 68 percent, 117 Johns words accounted for 67 percent, and 120 Hillerich words accounted for 67 percent.

Levels 2, 3, 4 and 5 are the preprimer, primer, and first grade readers in the Ginn (1979) 720 reading series ("Canadianized" version). They contained 11,148 tokens and 549 types (T/T=.05). Sixty-four percent of the vocabulary in the four levels was represented by 139 I.C. words. This compares with the other lists: 1) 139 Dolch words represented 64 percent of the vocabulary, 2) 138 Johns words represented 63 percent of the vocabulary, 3) 133 Durr words represented 65 percent of the vocabulary and 142 Hillerich words represented 60 percent of the vocabulary.

Allyn and Bacon (1977) Pathfinder Series levels 6, 7, 8 and 9 are the first grade readers and contain 11,531 words in connected discourse (745 types, T/T=.06). The I.C. list shares 171 words with the Allyn and Bacon which is 66 percent of the total vocabulary. The Dolch list accounted for 63 percent of the vocabulary in 163 words while the Johns list accounted for 61 percent in 153 words. One hundred fifty-six of the Durr words accounted for 64 percent of the Allyn and Bacon first grade vocabulary. Finally, 163 Hillerich words accounted for 59 percent of the Allyn and Bacon first grade vocabulary. Correlation coefficients were also calculated for the association between I.C. frequencies and those generated by the three reading series. A correlation coefficient of .84 was computed for Macmillan, .81 for Ginn, and .80 for Allyn and Bacon. There is a high degree of association between the frequencies of the words on the I.C. list and in the initial reading texts.

The I.C. list came from a small corpus of words, yet it was as representative of the cumulative vocabulary of the initial reading texts sampled as the other basic lists. The similarities were compelling. An intriguing question was prompted by these findings; would an even smaller sample produce a representative list?

### *The Second Grade List*

The second grade (S.G.) corpus included a total of 2,011 correctly spelled words with 361 different units, representing a mean of approximately 34 (33.50) words per essay. A rank order list of words representing the same proportion of the vocabulary as the I.C. list was compiled. Specifically prompted words were deleted and tied rankings were included. The S.G. list contains 105 words and is shown in Table 2.

Most of the S.G. words are on the lists discussed previously. Eighty-one of them are on the Dolch, 84 on the Johns, 82 on the Durr, 88 on the Hillerich, and 64 on the Thomas 100 most frequent list. An indication of the degree to which the second grade students' vocabulary is different from the older students in the study can be seen by the words on the S.G. list not on the I.C. list. They include such items as



*God, time, dad, and better.* Other items are clearly artifacts of the assignments, i.e., *box, room,* etc. Seventy-nine of the S.G. words represent 57 percent of the vocabulary in the Macmillan texts, 86 represent 54 percent of the Ginn vocabulary, and 95 represent 53 percent of the Allyn and Bacon vocabulary. The S.G. list accounts for over half of the vocabulary in the initial reading texts.

TABLE 2  
RANK ORDER LIST OF SECOND GRADE WORDS

|         |         |           |
|---------|---------|-----------|
| the     | that    | why       |
| is      | very    | am        |
| he      | want    | at        |
| and     | but     | lots      |
| you     | do      | no        |
| to      | have    | now       |
| I       | of      | said      |
| me      | took    | time      |
| box     | God     | when      |
| my      | all     | again     |
| it      | from    | anything  |
| a       | if      | be        |
| in      | on      | better    |
| she     | please  | birthday  |
| for     | we      | dad       |
| bad     | came    | eat       |
| man     | give    | end       |
| your    | good    | father    |
| so      | know    | going     |
| him     | one     | had       |
| not     | open    | how       |
| then    | tell    | into      |
| don't   | there   | just      |
| because | this    | make      |
| can     | too     | mom       |
| his     | went    | nice      |
| like    | food    | other     |
| saw     | gave    | put       |
| are     | her     | say       |
| with    | mother  | should    |
| room    | play    | something |
| will    | see     | they      |
| boy     | tall    | think     |
| us      | things  | what      |
| was     | walking | would     |

The correlations between the frequencies generated by the S.G. words and the large corpus lists are Durr .74, Rinsland .77, and Carroll, Davies, and Richman .64. The ordering of language appears to be a robust characteristic varying little from sample to sample. The unique items of the vocabulary of a particular population

are evident in relatively low frequency items. One more question was explored: does pooling the vocabulary of textbooks mask the uniqueness of individual texts?

The lowest level textbook from each series was analyzed. The Allyn and Bacon level six contained 1,836 words, the top ten representing 36 percent of the volume's vocabulary—*the, is, big, on, a, I, am, Ann, Nat, and Nate*. Forty-five percent of the series "r" level four vocabulary was represented by the top ten words—*I, the, and, why, to, jump, he, like, does, and says*. In descending order the ten words *Bill, and, Jill, Lad, is, can, here, this, Nan, and Ted* represented 55 percent of Ginn 720 level two. The T/T ratios for the three texts were, respectively, .07, .13, and .02. These initial texts varied in the rate of repetition of vocabulary (a low T/T indicates high repetition). The I.C. list accounted for 48 percent of the Ginn, 64 percent of the Macmillan, and 57 percent of the Allyn and Bacon vocabulary. Portions of the vocabulary accounted for by the other lists are: 1) Ginn—Dolch 44 percent, Durr 54 percent, Johns 54 percent, Hillerich 46 percent; 2) Macmillan—Dolch 71 percent, Durr 65 percent, Johns 64 percent, Hillerich 59 percent; 3) Allyn and Bacon—Dolch 50 percent, Durr 53 percent, Johns 54 percent, Hillerich 55 percent. Seventeen percent of the Macmillan vocabulary was not included on any of the lists (two percent were proper nouns). Thirty-five percent of the Ginn vocabulary did not occur on any of the lists (30 percent were proper nouns). Thirty percent of the Allyn and Bacon vocabulary was not found on any of the lists (17 percent were proper nouns). The S.G. list, with considerably fewer words accounted for 45 percent of the Macmillan, 43 percent of the Ginn, and 44 percent of the Allyn and Bacon vocabularies.

### Discussion

The I.C. word list accounts for a substantial portion of the vocabulary in the initial reading texts analyzed. The percentage of vocabulary the I.C. list represents is as high or higher in all cases as the percentage represented by the Dolch list, the Durr list, the Hillerich list, and the Johns RDL.

The similarity of the ordering of words between the I.C. and other vocabulary lists is compelling. The I.C. ordering is most similar to the Rinsland list which is also based on written material. It is also similar to the vocabulary in children's trade books, although to a lesser degree. It is least similar to the vocabulary of textbooks used in grades three to nine (Carroll, Davies, & Richman).

The similarity between the I.C. list and Thomas' Canadian list is compelling. Fifty-eight percent of all the words in the Thomas list were accounted for by one hundred items, the I.C. list containing 94 of them. The similarities between the vocabularies of Canadian students and inner-city American students are striking.

The I.C. list was developed from a relatively small corpus of words, yet it appears to be representative when compared with other lists and with the vocabulary of initial reading texts. A list was developed from an even smaller corpus of words, the second grade essays. The S.G. list of 105 words accounts for 53 to 57 percent of the vocabulary in the initial reading texts, remarkable considering the average size of the essays was approximately 34 words. It appears the relative frequencies of words (and particularly function words) remains fairly constant from sample to sample, supporting Zipf's contention that the distribution of word frequencies occurring in any *ad libitum* sample of language always has the same distribution (Zipf, 1935). He notes,

Perhaps the most interesting feature of this high degree of orderliness in the distribution of words in the stream of speech is this: we select and arrange our words according to their meanings with little or no conscious reference to the relative frequency of occurrence of those words in the stream of speech, yet we find that words thus selected and arranged have a frequency distribution of great orderliness which for a large portion of the curve seems to be constant for language in general. (p. 48)

The writing samples were typical classroom assignments and so were not *ad lib*, but when text-prompted items were removed the list became relatively representative. The vocabulary of individual initial reading texts is not *ad libitum*. Two types of texts appeared. The first was characterized by the use of names and repetition while the second was characterized by the use of pronouns and verbs. The word lists accounted for considerably more of the vocabulary in the second type of text than in the first. The use of names and repetition was a feature of the initial reading texts of the forties and fifties. Have we given Dick and Jane new names?

Word lists are successful in representing substantial portions of the vocabulary in reading texts when the vocabulary of the reading texts is analyzed as a combined sample. To ascertain the vocabulary load for a particular text, one must analyze that text. It appears that pedagogical lists are more accurate in preparing students for cumulative vocabulary than they are for the vocabulary of individual texts.

### Conclusion

Teachers use word lists as sources of words to be taught to students as sight words. Authors attempt to compile lists that represent large portions of the vocabulary encountered in school texts. The general belief is that students who are able to recognize important sight words will have command of the majority of words they encounter in text. While this does not guarantee they will comprehend the text, many teachers believe it will aid comprehension since less time and effort need to be directed to recognition activities.

Findings of the present study show that a word list can be developed from a small language sample of a small number of subjects. The I.C. word list is as successful in accounting for the cumulative vocabulary in initial reading texts as are other lists.

It was hypothesized that the I.C. list would account for much of the vocabulary of initial reading texts. This hypothesis was confirmed. There are many published word lists and, as shown in this paper, they represent considerable portions of the printed text students will cumulatively encounter. The I.C. list is as successful as several other well-known lists in representing the vocabulary of initial reading texts. Hillerich (1974) notes, "wars may be *escalated*, and people *fragged* or *televised* while few sit in a *parlor* anymore, but we have yet to replace or supplement *the*, *and*, or *to*." This would seem true for the written language of students in the inner-city where "standard English" is not the mode. Indeed, it would seem to be true for students in different regions across Canada and the United States.

Pedagogical lists are most often compiled by consulting other lists that are derived from large samples of words. As seen here, this is not necessary. Indeed, the unique character of the vocabulary of subjects and individual textbooks is lost when the sample size is large. This phenomenon is particularly compelling in the case of initial reading texts. As far as pedagogical lists are concerned, they all represent significant portions of the cumulative vocabulary in initial reading texts, somewhere between 60 and 70 percent. So any major pedagogical list will serve us equally well.



We should begin, however, to look at the vocabulary not represented by the major lists, the 30 to 40 percent of the vocabulary that is not included. To meet the regional needs of a particular area we must analyze the relatively low frequency items of vocabulary that represent its unique character. In this manner the distinctively Canadian vocabulary can be identified in various geographic areas. The availability of personal computers to the classroom makes such a task possible for teachers. As far as pedagogical lists are concerned, we have done enough.

### References

- Allyn & Bacon. (1977). *Pathfinder. Allyn and Bacon Reading Program*. Boston: Author.
- Carroll, J. B., Davies, P., & Richman, B. (1971). *Word frequency book*. Boston: Houghton Mifflin; New York: American Heritage.
- Dolch, E. W. (1936, February). A basic sight vocabulary. *Elementary School Journal*, 36, 456-460.
- Dolch, E. W. (1960). *Teaching primary reading*. Champaign, IL: Garrard.
- Durr, W. K. (1973). A computer study of high-frequency words in popular trade juveniles. *The Reading Teacher*, 27(14), 37-42.
- Ginn and Company. (1979). *Reading 720*. Scarborough, ON.
- Hillerich, R. L. (1974, January). Word lists—getting it all together. *The Reading Teacher*, 27(1), 353-360.
- Johns, J. L. (1974). Should the Dolch list be retired, replaced, or revised? *Elementary School Journal*, 74, 375-380.
- Johns, J. L. (1976a). Updating the Dolch basic sight vocabulary. *Reading Horizons*, 16, 104-111.
- Johns, J. L. (1976b). Some comparisons between the Dolch basic sight vocabulary and the word list for the 1970s. *Reading World*, 15(3), 144-150.
- Johnson, D. (1971a, February). The Dolch list reexamined. *The Reading Teacher*, 24(5), 449-457.
- Johnson, D. (1971b, October). A basic vocabulary for beginning reading. *Elementary School Journal*, 72, 29-34.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Lowe, A. J., & Follman, J. (1974, October). Comparisons of the Dolch list with other word lists. *The Reading Teacher*, 28(1), 40-44.
- Macmillan Publishing Company. (1980). *Series r, Macmillan reading*. New York: Author.
- Mangieri, J. N., & Kahn, M. S. (1977, March). Is the Dolch list of 220 basic sight words irrelevant? *The Reading Teacher*, 30(6), 649-651.
- Mangieri, J. N. (1978, October). Dolch list revisited. *Reading World*, 18(1), 91-95.
- Miller, A. (1975). *A word counting and frequency analysis program*. Vancouver, BC: University of British Columbia Computing Centre.
- Murphy, H. A. (1957, December). The spontaneous speaking vocabulary of children in the primary grades. *Journal of Education*, 140(2), 2-105.
- Rinsland, H. D. (1947). *A basic vocabulary of elementary school children*. New York: The Macmillan Co.
- Thomas, V. (1972, December). The basic writing vocabulary of elementary school children. *The Alberta Journal of Educational Research*, 18(4), 243-248.
- Thomas, V. (1974). *Teaching spelling Canadian word lists and instructional techniques*. Toronto: Gage.

Zipf, G. K. (1965). *The psycho-biology of language, an introduction to dynamic philology*. Cambridge, MA: M.I.T. Press (original work published in 1935).

## APPENDIX

### *Description Prompt:*

We all know someone who is special. People can be special in many different ways. Describe a person who is special to you. Tell why you think he or she is special. If you do not know how to spell a word, just spell it the way you think it sounds. Since we do not care about misspelled words, please do not ask for help in spelling.

### *Argument Prompt:*

Pretend your best friend has just started smoking for the first time and you have decided to write him a letter. Since you know smoking is not good for your friend, you must try to get him to stop. You are going to do this by writing him a letter. If you do not know how to spell a word, just spell it the way it sounds. Since we do not care about misspelled words, please do not ask for help in spelling.

### *Narration Prompt:*

Look carefully at the three pictures. Please write a story about these three pictures. Tell what is going on in each picture. What do you think is happening? Remember, the three pictures go together so make one story. If you do not know how to spell a word, just spell it the way you think it sounds. Since we do not care about misspelled words please do not ask for help in spelling.



STEPHEN T. CAREY

*University of Alberta*

and

JAMES CUMMINS

*Ontario Institute for Studies in Education*

## Communication Skills in Immersion Programs

*A study of cognitive and behavioral variables, academic achievement, and English and French speaking skills was undertaken to find the correlates of high French oral communication skills of grade three francophone and anglophone immersion students as rated by teachers and independent francophone raters. Teachers' ratings correlated with academic achievement and cooperativeness variables whereas those of independent francophone raters did not. The need for further research and the development of objective procedures for the evaluation of oral communications skills is discussed in terms of the influence that such evaluations might have on students' attitudes and hence the success of immersion programs.*

One of the concerns related to the evaluation of immersion programs is the lack of adequate assessment of oral production and aural comprehension skills that have been performed for these programs. This deficiency exists in part due to the time constraints and monetary restrictions that exist for most evaluation programs, as well as the difficulty in finding qualified oral examiners.

Nevertheless, since one of the prime goals of immersion programs is to develop functional oral communication skills, it is important that evaluation techniques be developed, refined and understood in terms of their ecological validity. The present study sought to investigate what variables correlated with teachers' ratings of their students' oral French skills and how these correlations between variables and ratings compared with similar ratings by independent francophone teachers. A com-

---

Stephen T. Carey obtained his doctorate from the University of Toronto and teaches at the Faculté Saint-Jean of the University of Alberta. His research interests include memory and bilingualism.

James Cummins obtained his doctorate from the University of Alberta and teaches at the Modern Language Centre of the Ontario Institute for Studies in Education. His research interests center on cognitive aspects of bilingualism and language acquisition.



parison of these sets of ratings, it was believed, might shed light on the process of evaluation of students' oral skills by their teachers.

An additional concern of researchers, teachers, and parents is to identify the cognitive, behavioral, and background variables that are associated with students who do well in immersion programs, in order that some objective basis might be established for predicting if a given student will or will not do well in an immersion program. Such information would be useful in counseling parents with regard to placing or withdrawing their children from an immersion program. The present study, therefore, sought to determine what cognitive, behavioral, and background variables are associated with students who do well in immersion programs to gain further insight into the dynamics of these programs.

Students in grade three were examined in this study rather than children in the later grades for several reasons. First, the authors believe that students' initial progress in acquiring French and English communication skills is instrumental in determining their attitude towards the immersion program. Second, the authors believe that the early establishment of a positive attitude towards the French immersion program is instrumental to later success in the program. Thus, it is suggested that the students' perception of their initial success may be critical in determining their subsequent achievement in this domain, and inasmuch as the child's perceptions of his success is influenced by the teacher's perception of his competence, a further understanding of the basis of teacher's evaluations is critical.

Language achievement testing at this level is problematic and the authors chose to rely on francophone teachers and independent francophone testers for providing criterion measures of students' oral French achievement. In addition, a study of the relationship between teachers' ratings of francophone and anglophone students' oral French achievement and the ratings of independent francophones of the same students, it was believed, might provide insights into how teachers perceived each of these student groups. Finally, the authors believe that a study of the relationship between cognitive, behavioral, and background variables of the students and the evaluation of their French achievement by the two rating groups might reveal differences in teachers' perception of the students and also lead to a description of students who are likely to achieve well in immersion programs.

### *Subjects*

The subjects were 43 children drawn from all of the grade three classes of three Edmonton Catholic Separate School (ECSS) bilingual schools. Teachers were asked to specify the linguistic background of their pupils as either anglophone or francophone. Each child was administered the Raven's Progressive Matrices (a nonverbal intelligence test). The 43 children in the study were selected on the basis of having anglophone and francophone groups who were equivalent in terms of Raven's scores. However, when the children were classified according to their responses on the linguistic background questionnaire (Appendix C), it was found that many of those students classified by teachers as coming from English backgrounds, actually came from homes where a considerable amount of French was spoken (at least by parents to children). Consequently, the groups were classified into English, French, and mixed French-English home/backgrounds. According to this classification, 24 children came from English homes, 10 from predominantly French and 9 from mixed French-English homes.

### *Instruments*

The Teacher Rating and students' questionnaires were the same as those used in the Carey and Cummins (1983) grade five study and are presented in Appendices B and C. Based on children's responses to items 1-4 on the student questionnaire, an index of overall home French usage was constructed where 1 indicated English only spoken, 2 indicated some French spoken, 3, a considerable amount of French spoken, and 4, French only spoken. Groups three and four formed the francophone group and group two the mixed French-English group. A rating of two signified that at least one parent spoke French most often to the child while a rating of three signified that either both parents spoke French most often to the child or that the child interacted most often in French with at least one parent (i.e., both parent and child speak mostly French to one another). This variable is number 15 in Table 1.

English achievement was assessed by means of the Gates McGinitie Reading Test and Vocabulary and Comprehension Scales, administered by the ECSS at the grade two level. A measure of *English* syntax comprehension developed by Tremaine (1975) was also administered. The Primary Mental Abilities (PMA) IQ test, administered by the ECSS at the grade one level, was also included in the analysis.

French skills were assessed in three ways: first, by means of a *French* syntax comprehension test developed by Tremaine (1975). Five syntactical constructions were assessed for both French and English syntax comprehension. English and French versions were translations of one another, and the French version was administered first to all subjects (see Appendix A for examples of test items). Approximately two weeks separated administration of English and French versions. Each of the sentences was read to the subjects who were required to pick one of three pictures which illustrated the correct meaning of the sentence. Four sentence exemplars were read for each of the five syntax construction types. Order of sentence presentation was random. The five constructions were as follows (from easiest to most difficult) according to the norms supplied by Tremaine (1975).

1. *Relative Clause*  
The baby eats the cake which is on the table.  
The baby on the table eats the cake.
2. *Passive*  
The baby is seen by the girl.  
The girl is seen by the baby.
3. *Reflexive*  
The boys see themselves.  
The boys see each other.
4. *Direct Object*  
The girl shows the cow to the dog.  
The girl shows the dog to the cow.
5. *For-to*  
The baby gives the ball for the dog to the cat.  
The baby gives the ball for the cat to the dog.

Teachers were also asked to rate children's proficiency in French on four five-point scales: French reading vocabulary, French reading comprehension, French oral expression, and French aural comprehension. These four measures correlated very highly (range .65-.94) with each other indicating that teachers were possibly in-

adventently giving a global rating of French proficiency rather than discriminating between different subskills. Consequently, a combined score from the four scales was used in the analysis.

TABLE 1

| Variable |                                       | Teacher<br>French Rating | French<br>Speaking Rating<br>(Testers) |
|----------|---------------------------------------|--------------------------|--|
| 1.       | Teacher French Rating                 | 1.00                     | .23                                    |
| 2.       | French Speaking Rating                | .23                      | 1.23                                   |
| 3.       | Gates Vocabulary                      | .44*                     | -.41*                                  |
| 4.       | Gates Comprehension                   | .51*                     | -.23                                   |
| 5.       | Raven's                               | .41*                     | .16                                    |
| 6.       | PMA                                   | .39*                     | -.29                                   |
| 7.       | Language Spoken by Mother (SQ1)       | -.08                     | .47*                                   |
| 8.       | Language Spoken by Father (SQ2)       | .22                      | .22                                    |
| 9.       | Language Spoken to Mother (SQ3)       | .23                      | .27                                    |
| 10.      | Language Spoken to Father (SQ4)       | .22                      | .07                                    |
| 11.      | Language Spoken to Siblings (SQ5)     | .07                      | .05                                    |
| 12.      | Happy to Speak French Elsewhere (SQ6) | .51*                     | .09                                    |
| 13.      | Language Preference to Teacher (SQ7)  | .51*                     | .07                                    |
| 14.      | Language Preference to Friends (SQ8)  | .06                      | -.10                                   |
| 15.      | Overall Home French Usage             | .19                      | .37*                                   |
| 16.      | English Syntax (Es) Total             | .38*                     | .20                                    |
| 17.      | Es Relative                           | .33*                     | .08                                    |
| 18.      | Es Passive                            | .19                      | -.32*                                  |
| 19.      | Es Referential                        | .29                      | -.02                                   |
| 20.      | Es Direct Object                      | .28                      | .36*                                   |
| 21.      | Es For-To                             | .15                      | .04                                    |
| 22.      | French Syntax (Fs) Total              | .54*                     | .17                                    |
| 23.      | Fs Relative                           | .30*                     | .06                                    |
| 24.      | Fs Passive                            | .40*                     | -.04                                   |
| 25.      | Fs Referential                        | .24                      | -.19                                   |
| 26.      | Fs Direct Object                      | .09                      | -.04                                   |
| 27.      | Fs For-To                             | .45*                     | .35*                                   |
| 28.      | Asks Questions (TR1)                  | .41*                     | .03                                    |
| 29.      | Responds to Questions (TR2)           | .61*                     | .14                                    |
| 30.      | Group Participation (TR3)             | .58*                     | .10                                    |
| 31.      | Friends (TR4)                         | .59*                     | -.14                                   |
| 32.      | Lively (TR5)                          | .52*                     | -.24                                   |
| 33.      | French at Recess (TR6)                | .61*                     | .20                                    |
| 34.      | Asks Teacher Help (TR7)               | .26                      | -.22                                   |
| 35.      | French in Corridors (TR8)             | .56*                     | .19                                    |
| 36.      | Extra French at Home (TR9)            | .46*                     | .24                                    |
| 37.      | Sex (2= male, 1= female)              | .35*                     | .23                                    |

\* =  $p \leq .05$

\*\* =  $p \leq .01$



| Variable |                                       | Gates<br>Vocabulary | Gates<br>Comprehension |
|----------|---------------------------------------|---------------------|------------------------|
| 1.       | Teacher French Rating                 | .44*                | .51*                   |
| 2.       | French Speaking Rating                | -.41*               | -.23                   |
| 3.       | Gates Vocabulary                      | 1.00                | .81                    |
| 4.       | Gates Comprehension                   | .81*                | 1.00                   |
| 5.       | Raven's                               | .21                 | .40*                   |
| 6.       | PMA                                   | .71*                | .69*                   |
| 7.       | Language Spoken by Mother (SQ1)       | -.47*               | -.53*                  |
| 8.       | Language Spoken by Father (SQ2)       | -.19                | -.22                   |
| 9.       | Language Spoken to Mother (SQ3)       | -.16                | -.24                   |
| 10.      | Language Spoken to Father (SQ4)       | -.01                | -.06                   |
| 11.      | Language Spoken to Siblings (SQ5)     | -.22                | -.19                   |
| 12.      | Happy to Speak French Elsewhere (SQ6) | -.16                | -.34*                  |
| 13.      | Language Preference to Teacher (SQ7)  | .38*                | .41*                   |
| 14.      | Language Preference to Friends (SQ8)  | .17                 | .03                    |
| 15.      | Overall Home French Usage             | -.23                | -.30*                  |
| 16.      | English Syntax (Es) Total             | .35*                | .51*                   |
| 17.      | Es Relative                           | .44*                | .50*                   |
| 18.      | Es Passive                            | .38*                | .43*                   |
| 19.      | Es Referential                        | .13                 | .17                    |
| 20.      | Es Direct Object                      | .13                 | .34*                   |
| 21.      | Es For-To                             | .14                 | .11                    |
| 22.      | French Syntax (Fs) Total              | .37*                | .48*                   |
| 23.      | Fs Relative                           | .27                 | .12                    |
| 24.      | Fs Passive                            | .26                 | .45*                   |
| 25.      | Fs Referential                        | .16                 | .12                    |
| 26.      | Fs Direct Object                      | .24                 | .42*                   |
| 27.      | Fs For-To                             | .29                 | .27                    |
| 28.      | Asks Questions (TR1)                  | .31*                | .39*                   |
| 29.      | Responds to Questions (TR2)           | .30*                | .44*                   |
| 30.      | Group Participation (TR3)             | .27                 | .46*                   |
| 31.      | Friends (TR4)                         | .38*                | .54*                   |
| 32.      | Lively (TR5)                          | .39*                | .48*                   |
| 33.      | French at Recess (TR6)                | .32                 | .41*                   |
| 34.      | Asks Teacher Help (TR7)               | .10                 | .26                    |
| 35.      | French in Corridors (TR8)             | .10                 | .27                    |
| 36.      | Extra French at Home (TR9)            | .11                 | .17                    |
| 37.      | Sex (2= male, 1= female)              | .19                 | .10                    |

\* =  $p \leq .05$ \*\* =  $p \leq .01$ 

The third measure of French proficiency was a rating of French speech production made by a native francophone teacher on the basis of a 10-15 minute interview with each child. The interview consisted of casual conversation related to the child's interests, as well as activities such as describing a set of cartoon pictures and listening to and retelling a French story. The rating was made on a 7-point scale, ranging from "speaks French like a native speaker of his/her age" to "has almost no command of spoken French."

*Results*

The correlations of the variables with French and English language skills are presented in Tables 1 and 2 and differences between the three groups of students are presented in Table 3.

TABLE 2

| Variable |                                       | French<br>Syntax Total | English<br>Syntax Total |
|----------|---------------------------------------|------------------------|-------------------------|
| 1.       | Teacher French Rating                 | .54*                   | .38*                    |
| 2.       | French Speaking Rating                | .17                    | .20                     |
| 3.       | Gates Vocabulary                      | .37*                   | .35*                    |
| 4.       | Gates Comprehension                   | .48*                   | .51*                    |
| 5.       | Raven's                               | .37*                   | .34*                    |
| 6.       | PMA                                   | .51*                   | .48*                    |
| 7.       | Language Spoken by Mother (SQ1)       | .07                    | -.13                    |
| 8.       | Language Spoken by Father (SQ2)       | -.08                   | -.06                    |
| 9.       | Language Spoken to Mother (SQ3)       | .25                    | -.05                    |
| 10.      | Language Spoken to Father (SQ4)       | .04                    | -.03                    |
| 11.      | Language Spoken to Siblings (SQ5)     | .20                    | -.01                    |
| 12.      | Happy to Speak French Elsewhere (SQ6) | .27                    | .36*                    |
| 13.      | Language Preference to Teacher (SQ7)  | .57*                   | .25                     |
| 14.      | Language Preference to Friends (SQ8)  | .09                    | .13                     |
| 15.      | Overall Home French Usage             | .12                    | -.08                    |
| 16.      | English Syntax (Es) Total             | .66*                   | 1.00                    |
| 17.      | Es Relative                           | .37*                   | .64*                    |
| 18.      | Es Passive                            | .24                    | .40*                    |
| 19.      | Es Referential                        | .33*                   | .50*                    |
| 20.      | Es Direct Object                      | .49*                   | .65*                    |
| 21.      | Es For-To                             | .33*                   | .47*                    |
| 22.      | French Syntax (Fs) Total              | 1.00                   | .66*                    |
| 23.      | Fs Relative                           | .43*                   | .19                     |
| 24.      | Fs Passive                            | .50*                   | .35*                    |
| 25.      | Fs Referential                        | .49*                   | .16                     |
| 26.      | Fs Direct Object                      | .47*                   | .52*                    |
| 27.      | Fs For-To                             | .70*                   | .47*                    |
| 28.      | Asks Questions (TR1)                  | .49*                   | .28                     |
| 29.      | Responds to Questions (TR2)           | .54*                   | .37*                    |
| 30.      | Group Participation (TR3)             | .48*                   | .33*                    |
| 31.      | Friends (TR4)                         | .25                    | .18                     |
| 32.      | Lively (TR5)                          | .19                    | .11                     |
| 33.      | French at Recess (TR6)                | .46*                   | .30*                    |
| 34.      | Asks Teacher Help (TR7)               | .25                    | .07                     |
| 35.      | French in Corridors (TR8)             | .49*                   | .30*                    |
| 36.      | Extra French at Home (TR9)            | .38*                   | .29                     |
| 37.      | Sex (2= male, 1= female)              | .30*                   | .07                     |

\* =  $p \leq .05$

\*\* =  $p \leq .01$

TABLE 3  
T-TESTS FOR DIFFERENCES BETWEEN GRADE 3 GROUPS

|     | Anglophone | (N=24) | Means<br>Francophone | (N=10) | Mixed | (N=9) |
|-----|------------|--------|----------------------|--------|-------|-------|
| 1.  | 14.3       |        | 16.00                |        | 12.8  |       |
| 2.  | 41.8       | **     | 51.0                 | **     | 36.1  |       |
| 3.  | 36.7       |        | 29.1                 |        | 38.3  |       |
| 4.  | 39.4       | **     | 28.1                 |        | 33.5  |       |
| 5.  | 28.6       |        | 29.6                 |        | 25.6  |       |
| 6.  | 114.4      |        | 108.2                |        | 112.8 |       |
| 7.  | 2.0        | **     | 3.8                  | **     | 2.7   |       |
| 8.  | 2.0        | **     | 3.6                  | **     | 2.3   |       |
| 9.  | 2.0        | **     | 3.4                  | *      | 2.4   |       |
| 10. | 2.0        | **     | 3.1                  | *      | 2.2   |       |
| 11. | 2.0        | **     | 2.7                  |        | 2.3   |       |
| 12. | 3.0        | *      | 3.6                  | *      | 2.6   |       |
| 13. | 2.5        |        | 2.6                  |        | 2.6   |       |
| 14. | 1.4        |        | 1.6                  |        | 1.4   |       |
| 15. | ---        |        | ---                  |        | ---   |       |
| 16. | 13.3       |        | 12.5                 |        | 13.0  |       |
| 17. | 3.3        |        | 3.0                  |        | 3.1   |       |
| 18. | 3.3        | **     | 2.4                  | *      | 3.3   |       |
| 19. | 2.5        |        | 2.4                  |        | 2.8   |       |
| 20. | 3.0        | *      | 2.8                  |        | 2.3   |       |
| 21. | 1.2        |        | 1.9                  |        | 1.4   |       |
| 22. | 13.8       |        | 14.0                 |        | 14.0  |       |
| 23. | 3.2        |        | 3.5                  |        | 3.8   |       |
| 24. | 3.4        |        | 3.3                  |        | 3.3   |       |
| 25. | 2.5        |        | 2.0                  |        | 2.7   |       |
| 26. | 3.1        |        | 2.7                  |        | 3.0   |       |
| 27. | 1.6        |        | 2.4                  |        | 1.2   |       |
| 28. | 4.0        |        | 3.8                  |        | 4.4   |       |
| 29. | 4.7        |        | 4.0                  |        | 4.1   |       |
| 30. | 4.6        |        | 4.3                  |        | 4.4   |       |
| 31. | 4.5        |        | 4.0                  |        | 4.1   |       |
| 32. | 4.5        |        | 4.3                  |        | 4.3   |       |
| 33. | 2.3        |        | 3.1                  |        | 2.0   |       |
| 34. | 3.8        |        | 4.0                  |        | 4.1   |       |
| 35. | 2.1        | *      | 2.9                  |        | 2.0   |       |
| 36. | ---        |        | ---                  |        | ---   |       |

\* =  $p \leq .05$

\*\* =  $p \leq .01$

*Correlates of French Skills*

It is clear from the pattern of correlations that Teacher French Rating (TFR) and French Syntax Total (FST) are indicative of very different dimensions of proficiency as compared to French Speaking Rating by an independent rater (FSR). Tables 1 and 2 show that TFR and FST both correlate highly with measures of academic achievement or potential (e.g., Gates Vocabulary and Comprehension,



Raven's, PMA, English Syntax). Thus, TFR and FST together with these English measures appear to be assessing a dimension of academic proficiency which is independent of language of testing, language spoken at home, or French speech rating by an independent francophone rater. Thus, teachers rate students who score high on cooperativeness and academic success and English language skills as also being proficient in French. In contrast, native francophone raters' judgments of the students' oral French proficiency did not show these correlations.

Teacher French ratings were significantly related to most of the indices of teachers' ratings of students' behavior (variables 28-36). These variables appear to be tapping a dimension of academic cooperativeness. Thus, these relationships are consistent with our interpretation of the teachers rating academic and English competence, cooperativeness, and French as a unitary dimension. The strong relationships between TFR and FST and children's language preference to teacher (variable 13) is also consistent with this pattern.

French Speech Ratings (FSR) by independent raters present a different picture. They are positively related to the use of French at home, especially language spoken by the child's mother (variable 7) but tend to be less highly correlated with English academic skills than do teachers' French ratings. This latter finding can be accounted for by the fact that the francophone group tends to do more poorly on measures of English academic and cognitive skills (variable 4) than the anglophone group (see Table 3). The lack of significant relationships between FSR and intelligence and academic French skills is consistent with Genesee's (1976) finding that the development of communication skills in French immersion programs is independent of children's intellectual ability. These results also lend support to the proposed distinction between Basic Interpersonal Communication Skills (BICS) and Cognitive Academic Language Proficiency (CALP) (Cummins, 1979; Swain, 1981).

### *Correlates of English Skills*

The patterns of correlations for English academic skills are almost identical to those for French academic skills—compare, for example, the correlations for English and French syntax and for Gates Comprehension and Teacher French Rating. The only group of variables which do not conform to this pattern are the indices of home language use and children's willingness to speak French (variable 12) which are negatively related to Gates English Comprehension probably because of the lower scores of the francophone group on this variable.

### *Group Differences*

Because of the small numbers of students in the francophone and mixed groups, group differences must be interpreted cautiously. However, there are some interesting trends in the data. Since the home language background questionnaire was the basis for group selection, it is not surprising that there are group differences on variables 7-11. The francophone group also tends to be more willing to use French outside of school (variable 12) but there are no differences in French preference to teachers and friends among francophone, anglophone, and mixed groups (variables 13 and 14). The francophone group performed more poorly than the anglophone group on Gates Comprehension, but better than the anglophone and mixed groups on French Speaking Rating. They performed more poorly than both the other groups in their comprehension of passives in the English Syntax Test (variable 18) while the mixed group performed worse than the anglophone group in the under-

standing of direct and indirect objects (variable 20). The only other difference was that the francophone group was rated by teacher to use more French in the corridors (variable 35) as compared to the anglophone group.

### Discussion

The fact that the francophone group did not perform as well as the anglophone group on Gates Comprehension suggests that at the time when the test was administered (grade two), the francophone group had not yet caught up with the anglophone group in English academic skills. However, other data suggests that this has been accomplished by the grade five level (Carey & Cummins, 1983).

Two additional aspects of the findings seem worthy of special attention. First is the unitary nature of academic skills regardless of the specific language (English or French) of instruction; this means that students who do well in English programs will also do well in French immersion programs and that students who are not performing well in French immersion programs will not necessarily improve if they are transferred into a regular program. Secondly, the acquisition of French speaking skills in an immersion context as rated by independent native French speaking teachers, appears not to be dependent on general intellectual or academic ability. These findings are consistent with findings of previous research on French immersion programs. Genesee (1976) states that the acquisition of French literacy skills is related to the student's IQ level, but the acquisition of interpersonal communication skills is not. Similarly, Carey and Cummins (1983) reported that written performance on French cloze tests correlated highly with measures of intellectual and English academic skills, but not with oral skills.

The consistency of the results of the present study in conjunction with the pattern observed in previous research, warrants examining potential implications for predicting students' academic success. The findings suggest that the same factors underlie the development of academic skills in a French immersion program as in a regular English program. Thus, children who experience *academic* problems in a French program are likely to experience similar academic problems in an English program, and students who do well in unilingual English academic studies will also tend to do well in studies taught in French. A corollary of this is that academic problems will not necessarily disappear when a child who is experiencing difficulties in a French immersion program is switched to an English program. There is some evidence to support this hypothesis (Bruck, 1978, 1982). However, due to the paucity of research regarding the relationship between learning disabilities and success in immersion programs and the existence of contradictory data (Trites, 1976, 1978), no firm conclusions can be drawn at the present time.

The converse implication is that the development of oral communication skills appears to be independent of general academic and intellectual skills and that many children who experience academic difficulties in a French immersion program will nevertheless develop reasonable fluency in French. Future research should delineate the variables that are not associated with success in written French acquisition but are prerequisites for success in oral communication skills. Extroversion and social confidence would appear to be important assets in such immersion pedagogy since an emphasis is placed on oral production in a group setting. Indeed, there is some evidence (Bruck, 1982) that it is frequently the child who is reticent, shy and introverted that may experience difficulty in early immersion programs due to the requirements of oral production in a social group setting.



Finally, the inconsistencies in evaluating oral French competence between teachers and independent francophone raters is extreme in some cases. For example, the correlation between teachers' ratings of French and variable 3, Gates Vocabulary is +44, yet for independent French raters this correlation is -41. Similar reversals are found between the two rating groups for variable 4, Gates Comprehension, variable 5, Raven's Progressive Matrices and variable 6, Primary Mental Abilities. Clearly teachers' ratings of students' *French* achievement are highly correlated with English competence and general academic ability whereas the independent raters' evaluations show that proficiency in oral French communication is independent of such abilities. There is one major implication of this finding. The results demonstrate the need to include several independent measures of oral communication skills in any assessment research design. The issue of how to interpret the results is less clear cut. It appears that the teachers' assessment of the child's oral fluency is biased by their knowledge of the child's academic proficiency and the basic type of interaction they have with the child (i.e., whether the child is cooperative or non-cooperative), and this knowledge masks the teacher's ability to assess French oral expression as a discrete skill.

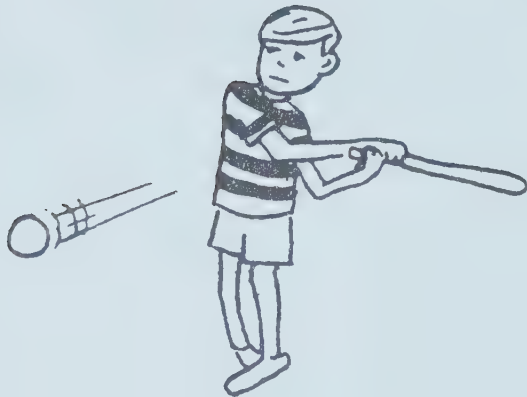
As previously mentioned, one of the major goals of immersion programs is to promote fluent oral communication skills. This research signals a caution in the teacher's ability to recognize this discrete skill and suggests the need for the availability of alternative objective methods, instruments and procedures for evaluating oral and written French skills in order to overcome the teacher's tendency to rate students' French competence in terms of their academic ability. This would seem to be particularly important for early grades in the immersion program since the student's self-perception and future attitude toward the immersion program are surely influenced by the teacher's evaluation of the student's French communication skills, as well as the teacher's perceptions of the student's language competence and his/her consequent behavioral correlates.

Finally, the implications for teacher preparation programs for immersion teachers would include a consideration of the critical role that students' attitudes play in determining their success in that program and the important role that teachers' evaluations must surely play in influencing the students' attitudes towards the immersion program. Clearly, the assessment of oral communication skills remains highly problematic (Canale, 1982); nevertheless the reliance on the communicative method of instruction in immersion programs requires that teachers have access to methodologies and procedures that can aid them in performing both formative and summative evaluation of oral communication skills. The shift in emphasis from language form to language use that is the hallmark of immersion programs aimed at producing oral communicative performance must be accompanied by an equal shift in the evaluation of communication skills.



Appendix A

English Syntax Test: Passives

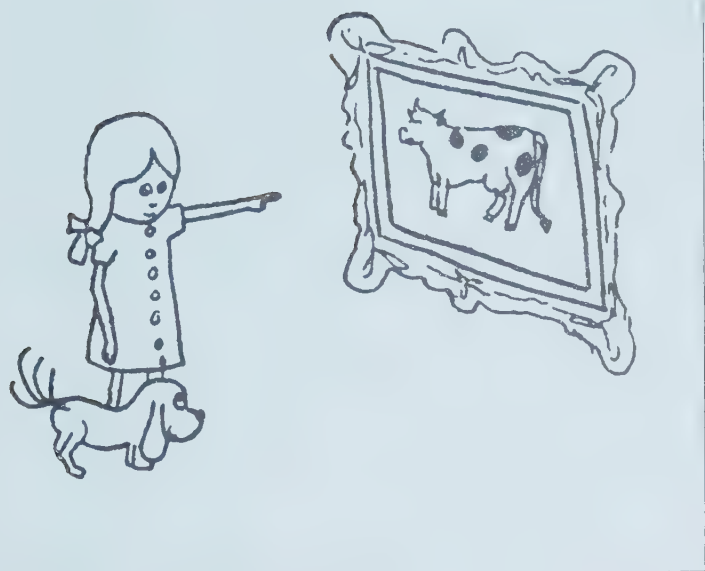


the boy hits the ball/  
the ball hits the boy

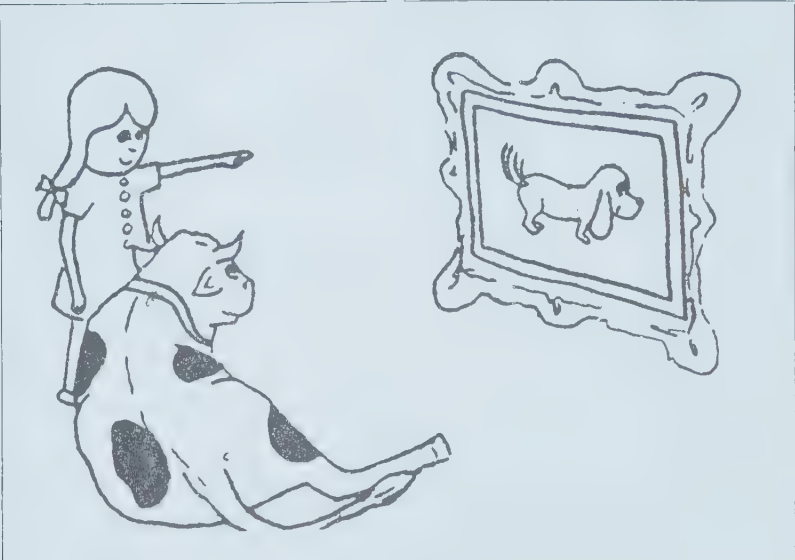
the boy is hit by the  
ball/the ball is hit  
by the boy



English Syntax Test: Direct/Indirect Object



the girl shows the  
cow to the dog/  
the girl shows the  
dog to the cow



Appendix B

Nom de l'Elève \_\_\_\_\_

En le comparant avec les autres enfants de votre classe et avec les autres enfants du même niveau que vous avez connus, est-ce que cet enfant:

|    |  |       |   |
|----|--|-------|---|
| 1. | Pose beaucoup de questions en classe.                                | _____ | Pose rarement de questions en classe.                             |
| 2. | Répond souvent aux questions posées en classe.                       | _____ | Répond rarement aux questions posées en classe.                   |
| 3. | Participe volontiers aux activités de groupe.                        | _____ | Ne participe pas aux activités de groupe.                         |
| 4. | A beaucoup d'amis.   | _____ | A très peu d'amis.  |
| 5. | Est volubile et plein de vie.  | _____ | Est tranquille, timide et à l'écart.                              |
| 6. | En récréation s'exprime toujours en français avec ses camarades.     | _____ | En récréation ne s'exprime jamais en français avec ses camarades. |
| 7. | Recherche l'aide et les conseils du professeur.                      | _____ | Travaille la plupart du temps seul/e.                             |
| 8. | S'exprime toujours en français dans les couloirs.                    | _____ | Ne s'exprime jamais en français dans les couloirs.                |
| 9. | Fait du travail supplémentaire en français à la maison ou à l'école. | _____ | Ne fait jamais du travail supplémentaire en français.             |

Appendix C

Name \_\_\_\_\_

School \_\_\_\_\_

- Which language does your *mother* speak at home most of the time?  
French \_\_\_\_\_ English \_\_\_\_\_ Other \_\_\_\_\_  
French and English about the same \_\_\_\_\_
- Which language does your *father* speak at home most of the time?  
French \_\_\_\_\_ English \_\_\_\_\_ Other \_\_\_\_\_  
French and English about the same \_\_\_\_\_
- Which language do *you* speak with your mother most of the time?  
Mostly French \_\_\_\_\_ Mostly English \_\_\_\_\_ Other \_\_\_\_\_  
French and English about the same \_\_\_\_\_
- Which language do *you* speak with your father most of the time?  
Mostly French \_\_\_\_\_ Mostly English \_\_\_\_\_ Other \_\_\_\_\_  
French and English about the same \_\_\_\_\_
- Which language do *you* speak with your brothers and sisters most of the time?  
Mostly French \_\_\_\_\_ Mostly English \_\_\_\_\_ Other \_\_\_\_\_  
French and English about the same \_\_\_\_\_
- How happy or unhappy are you to meet people outside of school whom you can speak French to?  
Very happy \_\_\_\_\_ Fairly happy \_\_\_\_\_ Fairly unhappy \_\_\_\_\_  
Very unhappy \_\_\_\_\_
- In school, do you prefer to speak English or French to your teacher?  
English \_\_\_\_\_ French \_\_\_\_\_ Both about the same \_\_\_\_\_
- In school, do you prefer to speak English or French to your friends?  
English \_\_\_\_\_ French \_\_\_\_\_ Both about the same \_\_\_\_\_



*References*

- Bruck, M. (1978). The suitability of early French immersion programs for the language disabled child. *Canadian Journal of Education*, 3, 51-72.
- Bruck, M. (1982). Language impaired children's performance in an additive bilingual education program. *Applied Psycholinguistics*, 3, 45-60.
- Canale, M. (1981). From communicative competence to communicative language pedagogy. In J. C. Richards & R. Schmidt (Eds.), *Language and communication*. London: Longman.
- Carey, S. T., & Cummins, J. (1983). Achievement, behavioral correlates and teachers' perceptions of francophone and anglophone immersion students. *Alberta Journal of Educational Research*, 29, 159-167.
- Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, 19, 197-205.
- Genesee, F. (1976). The role of intelligence in second language learning. *Language Learning*, 26, 267-280.
- Swain, M. (1981). Time and timing in bilingual education. *Language Learning*, 3, 1-16.
- Tremaine, R. V. (1975). *Syntax and Piagetian operational thought*. Washington, DC: Georgetown University Press.
- Trites, R. L. (1976). Children with learning difficulties in primary French immersion. *Canadian Modern Language Review*, 33, 193-216.
- Trites, R. L. (1978). Learning disabilities in immersion. *Canadian Modern Language Review*, 34, 888-889.

CLAIRE F. STAAB

*University of British Columbia*

## Student Teachers' Perceptions of Their Competence Versus the Perceptions of University Supervisors and Sponsor Teachers

*The primary purpose of this study was to determine whether student teachers, their university supervisors, and their sponsor teachers judged the student teacher's performance at the same level of competence over four clusters of teaching behaviors (i.e., Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities).*

*Sixty-six student teachers, their university supervisors, and their sponsor teachers were asked to respond to identical questionnaires, rating the student teacher's performance on the above behaviors. An item analysis using Lertap was done to determine the Hoyt estimate of reliability and each of the four clusters was subjected to a one way analysis of variance across three groups of subjects. Pearson correlations were also done to determine the degree of similarity of opinion both between and within groups. Results indicated a significant difference in Preparation for Instruction. Sponsor teachers rated student teachers significantly higher than university supervisors.*

Many articles over the past fifteen years have dealt with the notion that communication between student teachers, university supervisors and sponsor teachers needs to improve if desired changes are to be brought about in the student teacher's classroom performance (Larson & Ringness, 1965; Campbell & Williamson, 1973; Reynolds, 1977; Seiferth & Solliday, 1980; Queen & Mallen, 1982). However, little work has been done to determine whether the individual student teacher, his/her university supervisor, and his/her sponsor teacher view that student teacher's performance as being at the same level of competence and thus have a basis for beginning a constructive dialogue regarding student teacher performance. The purpose of this study, then, was to investigate whether in fact an individual student teacher,

---

Claire Staab (PhD, Arizona State) is an Assistant Professor at the University of British Columbia, teaching classes in developmental reading which have a student teaching component. She has been involved in student teaching programs for several years both at UBC and Arizona State.

his/her university supervisor and his/her sponsor teacher tend to judge that student teacher's performance as being at the same level of competence. In order to carry out the investigation, it was necessary to construct a reliable measure of student teacher performance, which may in itself prove useful after further investigation.

It would seem that the literature should have an abundance of studies indicating how the student teaching experience is viewed by each member of the student teaching triad (i.e., the student teacher, his/her university supervisor, and his/her sponsor teacher), and how the opinions of each member of the triad compare with one another. On the contrary, the present researcher, as well as Hattie, Olphert, and Cole (1982), in an extensive review of the literature, found a paucity of research on this question. The few studies which have been done comparing the views of members of the triad have produced mixed results. Joyce (1967) found that student teachers and supervisory personnel did not agree on a student's competence in teaching behaviors except in the area of discipline, or class control. Bueth and Pettibone (1973), however, found a relatively high correlation of agreement between sponsor teachers and student teachers when the two groups were asked to judge specific aspects of the student teacher's performance. Perhaps this high correlation in the Bueth and Pettibone study (1973) resulted from the fact that only highly specific teaching behaviors such as spelling errors and the student teacher's dependence upon lecture notes were examined. Using only a small sample of thirteen student teachers, Simmons (1973) found a high correlation between the opinions of university supervisors and sponsor teachers regarding such student teaching behaviors as lesson planning and lesson execution. Student teachers were not asked to rate their own level of competence. Mahan and Lester (1974) found a relatively high correlation between the responses of student teachers, sponsor teachers and university supervisors when members of the triad were asked to make judgments regarding their satisfaction with the general effectiveness of the student teacher program in regard to such issues as amount of professional preparation and rapport with other members of the triad; however, no specific teaching behaviors were assessed.

Although there have been few studies comparing the views of members of the triad regarding the student teacher's performance, several studies have indicated specific concerns of student teachers, without comparing these concerns with those of supervisory personnel. Freeland (1979) has listed the three major concerns of student teachers as follows: motivating distinterested pupils; handling discipline problems; and evaluating pupil progress. Bartos and Lotven (1980), in asking student teachers to rank fifteen competencies in terms of those most needed in the classroom, listed the following in order of importance: the ability to apply appropriate techniques, the ability to maintain order and assist students in developing self-discipline; and the ability to motivate students. Purcell and Seifert (1982) found student teachers to be most concerned with the problem of student discipline. None of the above studies assessed the opinions of the sponsor teacher or the university supervisor regarding what problems they felt the student teacher was experiencing or what behaviors were most essential.

As previously mentioned, communication between members of the triad is necessary if desired changes are to be brought about in the student teacher's classroom performance. This communication is, in fact, vital to the student teaching experience (Yee, 1969; Seiferth and Solliday, 1980; Tabachnick, Popkewitz, & Zeichner, 1980; Queen & Mallen, 1982). This communication must be one in which there is an



understanding of each other's point of view and a frank and positive discussion of mutual concerns (Rickard, 1982).

It seems highly unlikely that members of the triad will understand each other's point of view if there is not some basic agreement about the student teacher's level of competence and some understanding of what each member of the triad views as a problem or a particular strength in the student teacher's behavior. For example, it would be difficult for a university supervisor to discuss a need for improved lesson planning with a student teacher if that student teacher viewed his/her performance in this behavior to be adequate. In order for a positive discussion to take place, student teachers and supervisory personnel must recognize how they as individuals view teaching behaviors and how these behaviors are viewed by other members of the triad. Even though several studies have established the fact that communication between members of the triad is necessary, adequate research regarding this first step of initial agreement is still lacking. Do student teachers and their supervisory personnel tend to view a student teacher's behavior at the same level of competence and thus have a common basis for beginning necessary dialogue?

In the present study sixty-six student teachers, ten university supervisors, and fifty-seven sponsor teachers were asked to respond to a questionnaire, the development of which is described below. This questionnaire sought to gain information on four specific clusters of teaching behaviors—Preparation for Instruction, Use of Teaching Skills, Classroom Management and Personal Qualities, in order to answer the following research questions:

1. Do student teachers judge various aspects of their teaching performance as being at the same level of competence as do their university supervisors and their sponsor teachers?
2. When relationships within the four clusters (i.e., Preparation for Instruction, Use of Teaching Skills, Classroom Management and Personal Qualities) are compared, are the relationships among the three groups of subjects (i.e., student teachers, their university supervisors and their sponsor teachers) similar?
3. When relationships within single groups of subjects (e.g., student teachers) are examined, are the relationships among the four clusters (i.e., Preparation for Instruction, Use of Teaching Skills, Classroom Management and Personal Qualities) similar?

### *Method*

#### *Subjects*

Sixty-six student teachers, representing the entire population of a student teaching practicum, were asked to participate in the study. The practicum, comprising ten one-half day sessions in the classroom, was required for all elementary education students as an initial practicum and was conducted in the student teacher's second year of university study. Many of the student teachers had had previous experience in classrooms as observers, but this practicum was their first supervised teaching experience. The student teachers were assigned by the university student teaching office to an elementary school located within a large Western Canadian city or to a surrounding suburban area. The student teachers spent the ten sessions with one sponsor teacher in one classroom. The primary focus of this practicum was to prepare and teach lessons in reading, but the student teachers were also involved in helping the sponsor teacher with other subjects throughout the morning.

Fifty-seven sponsor teachers were asked to participate in the study. Some of these sponsor teachers supervised two student teachers simultaneously. Sponsor teachers were encouraged to conference with each student teacher weekly for at least one-half hour. These sponsor teachers comprised the total population of elementary school personnel supervising the student teachers.

Ten university supervisors participated in the study. These supervisors comprised the total population of university personnel supervising the student teachers. The university supervisors generally visited each student teacher in the classroom weekly and observed at least one entire lesson every two weeks. In most cases these university supervisors were also the student teacher's professor in a reading methods course which was taken simultaneously with the practicum. Student teachers also met with their university supervisors weekly for a one-hour seminar for the purpose of discussing problems encountered on the practicum, receiving feedback on lessons taught, and receiving additional help with teaching techniques. Individual conferences with the university supervisor were also scheduled as the student's need arose.

### *Procedure*

At the end of the ten week practicum, student teachers, sponsor teachers and university supervisors were asked to respond to an identical questionnaire containing 16 items. This questionnaire is labeled Table 1. The 16 items were grouped into four clusters of teaching behaviors (i.e., Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities). The respondents were asked to rate the student teacher's performance on these behaviors (i.e., separately the sponsor teacher and the university supervisor rated each of their own student teacher's performance and the student teacher rated his/her own performance). Each person responded to the questionnaire without any prior discussion with any other person. After data had been collected, each student teacher's responses were compared with those of his/her sponsor teacher and those of his/her university supervisor.

Neither the student teacher, the university supervisor, nor the sponsor teacher knew before the end of the practicum that they would be asked to fill out the questionnaire. To prevent collusion and to ensure that each questionnaire was answered independently by all respondents, certain precautions were taken. During the final week of the student teacher's experience the researcher visited the various student teaching seminars. The university supervisor conducting the seminar left the room and the student teachers responded to the questionnaires. These questionnaires were immediately collected by the researcher. Sixty-six student teachers were present in the seminars and each one responded to the questionnaire.

The following day the researcher asked each university supervisor to fill out the questionnaire on each student he/she supervised. These questionnaires were distributed and collected through university mail. All questionnaires distributed to university supervisors were returned.

The questionnaire was mailed to sponsor teachers during the week following the student teacher's last session. The sponsor teachers did not receive the questionnaire until contact with both the student teacher and the university supervisor had been terminated. A stamped, self-addressed envelope was provided for return to the researcher. All sponsor teachers except three returned the questionnaire.

All respondents were assured that their answers would not be available to anyone except the researcher and that their questionnaires would be coded to preserve anonymity.

TABLE 1  
STUDENT TEACHER PERFORMANCE EVALUATION

UNIVERSITY SUPERVISOR & SPONSOR TEACHER FORM

Please fill in your name below proper title and the name of the student teacher whom you are evaluating.

| Student Teacher   |       | University Supervisor  |       | Sponsor Teacher   |
|---|-------|--|-------|---|
| Directions: Put an 'X' in the parentheses beside the number which corresponds to your evaluation of the student teacher's performance in each category.<br>5 is excellent. 1 is unsatisfactory. |       |  |       |   |
| EXCELLENT   |       | SATISFACTORY   |       | UNSATISFACTORY  |
| A. PREPARATION FOR INSTRUCTION  |       |  |       |   |
| 5 ( )   | 4 ( ) | 3 ( )  | 2 ( ) | 1 ( )   |
| Lessons are consistently well prepared. Goal of lesson is thought through and teaching plans reflect this goal.   |       | Lessons are adequately prepared. There is some goal to the lesson and teaching plans indicate some connection with goal. |       | Problems with lessons plans. Goals seem confused. Teaching plans are sometimes vaguely related. |
| 5 ( )   | 4 ( ) | 3 ( )  | 2 ( ) | 1 ( )   |
| Appropriate teaching materials are consistently present during lesson.  |       | Appropriate materials are usually present during lesson.   |       | Problems in having appropriate materials for lesson.  |
| B. USE OF TEACHING SKILLS   |       |  |       |   |
| 5 ( )   | 4 ( ) | 3 ( )  | 2 ( ) | 1 ( )   |
| Lesson plan is consistently executed with appropriate pace.   |       | Lesson plan is usually executed with appropriate pace.   |       | Problems in executing lesson plan with appropriate pace.  |
| 5 ( )   | 4 ( ) | 3 ( )  | 2 ( ) | 1 ( )   |
| Lesson plan is consistently executed with appropriate sequence.   |       | Lesson plan is usually executed with appropriate sequence.   |       | Problems in executing lesson plan with appropriate sequence.                                    |
| 5 ( )   | 4 ( ) | 3 ( )  | 2 ( ) | 1 ( )   |
| Instructional strategies are consistently varied to fit lesson.   |       | Instructional strategies are usually varied to fit lesson.   |       | Problems varying instructional strategies to fit lesson.  |
| 5 ( )   | 4 ( ) | 3 ( )  | 2 ( ) | 1 ( )   |
| Examples and illustrations are consistently appropriate.  |       | Examples and illustrations are usually appropriate.  |       | Problems in using appropriate illustrations and examples.                                       |



Student Teachers' Perceptions of Competence

TABLE 1 (cont'd)

| EXCELLENT  |          | SATISFACTORY  |          | UNSATISFACTORY   |
|--|----------|---|----------|--|
| 5 (    )   | 4 (    ) | 3 (    )  | 2 (    ) | 1 (    )   |
| Consistently reinforces learners in ways which increase their motivation.                              |          | Usually reinforces learners in ways which increase their motivation.              |          | Problems reinforcing learners in ways which increase their motivation.                               |
| 5 (    )   | 4 (    ) | 3 (    )  | 2 (    ) | 1 (    )   |
| Lesson is consistently aimed at the ability of the learners.   |          | Lesson is usually aimed at the ability of the learners.                           |          | Problems aiming lesson at the ability of the learners.   |
| C. CLASSROOM MANAGEMENT  |          |   |          |  |
| 5 (    )   | 4 (    ) | 3 (    )  | 2 (    ) | 1 (    )   |
| Consistently anticipates difficulties and takes preventative measures.                                 |          | Usually anticipates difficulties and takes preventative measures.                 |          | Problems anticipating difficulties and taking preventative measures.                                 |
| 5 (    )   | 4 (    ) | 3 (    )  | 2 (    ) | 1 (    )   |
| Consistently firm but fair in response to misbehavior.   |          | Usually firm but fair in response to misbehavior.                                 |          | Problems being firm but fair in response to misbehavior.   |
| 5 (    )   | 4 (    ) | 3 (    )  | 2 (    ) | 1 (    )   |
| Consistently helps learners to take responsibility for their own actions and practice self-discipline. |          | Usually helps learners take responsibility and practice self-discipline.          |          | Problems in helping learners take responsibility for their own actions and practice self-discipline. |
| D. PERSONAL QUALITIES  |          |   |          |  |
| 5 (    )   | 4 (    ) | 3 (    )  | 2 (    ) | 1 (    )   |
| Consistently enthusiastic about teaching and communicates this enthusiasm to learners.                 |          | Usually enthusiastic about teaching and communicates this enthusiasm to learners. |          | Problems in being enthusiastic about teaching and communicating this enthusiasm to learners.         |
| 5 (    )   | 4 (    ) | 3 (    )  | 2 (    ) | 1 (    )   |
| Consistently maintains good rapport with school staff.   |          | Usually maintains good rapport with school staff.                                 |          | Problems maintaining good rapport with school staff.   |
| 5 (    )   | 4 (    ) | 3 (    )  | 2 (    ) | 1 (    )   |
| Consistently willing to accept constructive criticism and institute changes.                           |          | Usually willing to accept constructive criticism and institute changes.           |          | Problems receiving constructive criticism and instituting changes.                                   |

TABLE 1 (cont'd)

| EXCELLENT   |          | SATISFACTORY   |          | UNSATISFACTORY  |
|---|----------|--|----------|---|
| 5 (    )  | 4 (    ) | 3 (    )   | 2 (    ) | 1 (    )  |
| Voice projection is consistently adequate for the circumstances of the lesson (e.g., size of group, size of room, etc.) |          | Voice projection is usually adequate for the circumstances of the lesson (e.g., size of group, size of room, etc.) |          | Problems with voice projection.   |
| 5 (    )  | 4 (    ) | 3 (    )   | 2 (    ) | 1 (    )  |
| General appearance is consistently adequate to meet the school's standards of professionalism.                          |          | General appearance is usually adequate to meet the school's standards of professionalism.                          |          | Problems in maintaining a general appearance which is adequate to meet the school's standards of professionalism. |

Questionnaire Development

There were two slightly different formats of the questionnaire, one for the supervisors and sponsor teachers and one for the student teachers. The only difference in the formats was a change in verb form from third to first person (e.g., on the sponsor teacher and university supervisor's form, one item read, "Consistently helps students with self-discipline," while the student teacher's version read, "Consistently help students with self discipline." This slight change of verb form was necessary because student teachers were filling out the questionnaire in response to their own behavior and university supervisors and sponsor teachers were filling out the questionnaire in response to someone else's behavior (i.e., the student teacher's). The university supervisor and sponsor teacher form of the questionnaire appears as Table 1.

The 16 items on the questionnaire covered four main clusters of teaching performance and classroom behavior: Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities. These four clusters were obtained by reviewing the literature on necessary teaching skills (Peck & Tucker, 1973; Cooper, 1977; Chiarelott, Davidman, & Muse, 1980; Purcell & Seifert, 1982). Following this review of literature, twenty practicing elementary school teachers and five university professors involved in the supervision of student teachers were interviewed regarding the necessary teaching skills involved in student teaching. On the basis of this information and the literature review, a pilot questionnaire was prepared. This pilot questionnaire was given to eight university instructors and fifty-two student teachers. Neither these instructors nor these students were subjects in the final study. The respondents to the pilot questionnaire were asked to complete the instrument and also to indicate their opinion of its usefulness, its clarity of wording and format, and any problems in completing the form. The above procedures were necessary in establishing some estimate of content validity. The questionnaire which appears as Table 1 was developed as a result of this pilot questionnaire.

Analysis

A total of 50 questionnaires from each group (i.e., student teachers, university supervisors, and sponsor teachers) was analyzed. If either the student teacher, the

university supervisor, or the sponsor teacher did not complete the questionnaire or failed to complete all questions on the questionnaire, the data were not used.

Each of the four clusters which included Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities was treated separately. The results were analyzed as four clusters of teaching behaviors over three groups of subjects—student teachers, university supervisors, and sponsor teachers. An item analysis using Lertap (Nelson, 1974) was done to determine the Hoyt estimate of reliability, and each of the four clusters was subjected to a one-way analysis of variance across three groups of subjects. To determine the degree of similarity of opinion both between and within groups, Pearson correlations and tests of the significance of the difference between independently calculated correlations were calculated (Glass & Stanley, 1970).

TABLE 2  
MEANS, STANDARD DEVIATIONS, AND  
HOYT ESTIMATES OF RELIABILITY

| Groups                                  | Clusters of Teaching Behaviors     |                                     |                                     |                                     |
|---|------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
|   | Preparation for Instruction        | Use of Teaching Skills              | Classroom Management                | Personal Qualities                  |
| No. of individual items in each cluster | 2                                  | 6                                   | 3                                   | 5                                   |
| 1. Students<br>n=50                     | M=8.18<br>SD=(1.24)<br>Hoyt=(0.67) | M=22.28<br>SD=(2.66)<br>Hoyt=(0.60) | M=10.50<br>SD=(2.04)<br>Hoyt=(0.73) | M=22.18<br>SD=(2.44)<br>Hoyt=(0.66) |
| 2. University Supervisors<br>n=50       | M=7.86<br>SD=(2.13)<br>Hoyt=(0.93) | M=22.30<br>SD=(5.08)<br>Hoyt=(0.95) | M=10.62<br>SD=(2.65)<br>Hoyt=(0.93) | M=21.08<br>SD=(4.07)<br>Hoyt=(0.84) |
| 3. Sponsor Teachers<br>n=50             | M=8.78<br>SD=(1.50)<br>Hoyt=(0.88) | M=22.88<br>SD=(3.26)<br>Hoyt=(0.81) | M=10.30<br>SD=(2.22)<br>Hoyt=(0.83) | M=21.56<br>SD=(3.39)<br>Hoyt=(0.82) |

Results

Internal Consistency of Questionnaire

The Hoyt estimates of internal consistency of the four clusters of teaching behaviors for each of the three groups of subjects (i.e., student teachers, sponsor teachers, and university supervisors) are presented in Table 2. The estimate for the sponsor teachers and the university supervisors are within or above the levels usually accepted (i.e., .70 or greater) for group decisions (Cronback, 1972). However, only those of the university supervisors are above or close to the suggested reliability coefficient for individual decision making (i.e., .95) (Cronback, 1972). The reliability coefficients for the four clusters were somewhat lower for the student group



(.60 to .73). This finding may be a result either of the brevity of the four cluster scales, or a reflection of the degree to which beginning student teachers perceive their teaching behaviors, as unified clusters.

*Levels of Competence*

Table 2 presents the means and standard deviations for the four clusters of teaching behaviors across the three groups of raters.

Table 3 presents the results of the analyses of variance for the four clusters of teaching behaviors over the three groups of subjects.

TABLE 3  
ONE WAY ANALYSIS OF VARIANCE FOR FOUR CLUSTERS OF  
TEACHING BEHAVIORS AS RATED BY  
STUDENTS, UNIVERSITY SUPERVISORS AND SPONSOR TEACHERS

| Sources of Variation           | <i>df</i> | <i>MS</i> | <i>F</i> | <i>p</i> |
|--------------------------------|-----------|-----------|----------|----------|
| 1. Preparation for Instruction |           |           |          |          |
| Between Groups                 | 2         | 10.91     | 3.93     | 0.02     |
| Within Groups                  | 147       | 2.78      |          |          |
| 2. Use of Teaching Skills      |           |           |          |          |
| Between Groups                 | 2         | 5.80      | 0.40     | 0.67     |
| Within Groups                  | 147       | 14.50     |          |          |
| 3. Classroom Management        |           |           |          |          |
| Between Groups                 | 2         | 1.31      | 0.24     | 0.78     |
| Within Groups                  | 147       | 5.37      |          |          |
| 4. Personal Qualities          |           |           |          |          |
| Between Groups                 | 2         | 15.20     | 1.34     | 0.26     |
| Within Groups                  | 147       | 11.33     |          |          |

The only significant difference between the three groups was achieved in Cluster One: Preparation for Instruction. Multiple comparisons were done using the Scheffe test at the .05 level of significance. On Cluster One the student teachers rated themselves similarly to their sponsor teachers. The sponsor teachers, however, rated the student teachers significantly higher than the university supervisors.

*Inter-Cluster Relationships*

Table 4 presents the between-group correlations for each of the four clusters of items for each possible pair of subjects (e.g., Student Teachers and Sponsor Teachers). Of the twelve correlations, nine achieved statistical significance beyond the .01 level. The remaining three were not significant. Within each item cluster, tests of the significance of the difference between correlations revealed that no pair was different for Cluster One: Preparation for Teaching, and Cluster Four: Personal Qualities. For Cluster Three: Classroom Management, agreement between university supervisors and sponsor teachers was higher than that between student teachers and sponsor teachers ( $z=2.220, p<.05$ ), and that between student teachers and university supervisors ( $z=2.839, p<.01$ ). Although statistical significance was

not achieved on any pair of correlations for Cluster Two: Use of Teaching Skills, the difference in correlation between student teachers and university supervisors, and university supervisors and sponsor teachers was marginal ( $z=1.53$ ,  $p=0.126$ ).

TABLE 4  
BETWEEN GROUP PEARSON PRODUCT MOMENT CORRELATIONS OF  
STUDENT TEACHERS, UNIVERSITY SUPERVISORS, AND SPONSOR  
TEACHERS OVER FOUR CLUSTERS OF TEACHING BEHAVIORS  
( $n=50$ )

| Groups  | Clusters                            |                                |                              |                            |
|---|-------------------------------------|--------------------------------|------------------------------|----------------------------|
|   | 1<br>Preparation for<br>Instruction | 2<br>Use of Teaching<br>Skills | 3<br>Classroom<br>Management | 4<br>Personal<br>Qualities |
| Student Teachers to<br>University Supervisors | 0.52**                              | 0.28                           | -0.03                        | 0.45**                     |
| Student Teachers to<br>Sponsor Teachers       | 0.53**                              | 0.54**                         | 0.16                         | 0.45**                     |
| University Supervisors<br>to Sponsor Teachers | 0.41**                              | 0.64**                         | 0.58**                       | 0.56**                     |

\*  $p<.05$   
\*\*  $p<.01$

Table 5 presents the within-group correlations across the four clusters for each of the three groups of subjects. As may be seen, all but one of the 18 coefficients achieved statistical significance ( $p<.01$  in 16 cases and  $p<.05$  in one case). None of the differences between pairs of correlations within each group of subjects was statistically significant, thereby lending support to the view that each group of subjects perceived student teaching behaviors as being relatively unified or “global.” For example, if a student teacher was perceived as being competent in one behavior, he or she was perceived similarly on other behaviors.

Discussion and Implication

In general, it appears that the self perceptions of student teachers regarding their teaching performance are similar in mean levels of competence to the perceptions of their university supervisors and sponsor teachers. There was some difference found across groups of subjects in only one of the four clusters: Preparation for Instruction (i.e., the teaching behaviors of preparing lessons and having appropriate teaching materials available). Sponsor teachers tended to rate student teachers more positively on Preparation for Instruction than did the student’s university supervisor. One possible explanation for sponsor teachers rating higher than university supervisors is that universities in their student teacher programs normally require extensive and complete lesson plans which are then checked by the university supervisor before the student actually enters the classroom. Often these lesson plans are far more extensive than those which have been developed by the

sponsor teacher. University supervisors tend to consider the preparation of lesson plans as a vital element in the training of student teachers and thus are often more critical of these plans. It may be necessary for the university supervisor to take more time at the beginning of each practicum to discuss mutual expectations for lesson plans with the sponsor teacher.

Examination of Table 4 reveals that 8 of the 12 correlations between groups of subjects on the four cluster scores were statistically significant, despite the fact that each correlation accounted for only a moderate percentage of the total variance (18% to 40%). Agreement was not consistent across the three groups of subjects for the clusters, Preparation for Instruction and Personal Qualities. Thus, although university supervisors perceived student teachers as being at different levels of competence on Cluster One: Preparation for Instruction, they did agree on relative ranking. This finding supports the conjecture made above about the higher expectations of university supervisors with regard to lesson planning.

Less agreement among the three pair of subjects was found on Cluster Two: Use of Teaching Skills and particularly on Cluster Three: Classroom Management. It is interesting that there tends to be less agreement on the teaching behaviors which a student teacher must demonstrate during the course of a lesson in progress (i.e., classroom management and such teaching skills as lesson execution, using appropriate examples, and adjusting the planned lesson to the learner's needs) than in those teaching behaviors which a student teacher may initially bring to the lesson (i.e., a prepared lesson plan and various personal qualities).

As the within-group correlations in Table 5 reveal, it was also found that university supervisors and sponsor teachers tend to view student teachers as performing consistently. If student teachers receive a high rating in one particular teaching behavior then they are likely to receive a high rating in other teaching behaviors. Students are not viewed as weak in lesson planning, for example, and strong in classroom management. The reason for this finding is speculative at this time. One possible explanation is that the four teaching behaviors, Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities may be highly interdependent. For example, if a student teacher's management skills are poor and the class is out of control it is highly likely that the lesson will not be well executed, appropriate materials will not be used, and personal qualities will suffer. Conversely, if a student has not planned a lesson adequately, does not implement the lesson with the proper sequence and pace, it is also likely that students will not understand what is being presented and tend to cause problems for the student teacher in the area of classroom management. Thus a deficiency in one teaching behavior may radiate to other teaching behaviors. If these teaching behaviors are interdependent, university supervisors and sponsor teachers must be trained to view a student teacher's performance and begin to pinpoint behaviors which are causing general problems. For example, if a student teacher's performance is deficient because of a poorly planned lesson, then immediate attention must be given to lesson planning, although perhaps there were classroom management problems during the execution of the lesson.

A second possible explanation for the finding that university supervisors and sponsor teachers tended to view student teachers as performing consistently is that students simply did perform similarly in different areas of teaching behaviors. Clearly, there are many students who have a "feel for teaching" and simply do demonstrate ability in all teaching behaviors, and conversely, there are those students who are clearly lacking in most teaching behaviors. A third possible explanation



would be the operation of “halo” effect, or a tendency to judge seemingly diverse behavior at a single level of competence. Further research is needed before the possible reason or reasons can be anything but speculative.

TABLE 5  
WITHIN GROUP PEARSON PRODUCT MOMENT CORRELATIONS FOR  
STUDENT TEACHERS, UNIVERSITY SUPERVISORS AND SPONSOR  
TEACHERS OVER FOUR CLUSTERS OF TEACHING BEHAVIORS  
(n=50)  
Student Teachers

| Clusters | Clusters                            |                                |                              |                            |
|----------|-------------------------------------|--------------------------------|------------------------------|----------------------------|
|          | 1<br>Preparation for<br>Instruction | 2<br>Use of Teaching<br>Skills | 3<br>Classroom<br>Management | 4<br>Personal<br>Qualities |
| 1        |                                     | 0.59**                         | 0.22                         | 0.58**                     |
| 2        |                                     |                                | 0.57**                       | 0.62**                     |
| 3        |                                     |                                |                              | 0.33*                      |
| 4        |                                     |                                |                              |                            |

University Supervisors

| Clusters | Clusters                            |                                |                              |                            |
|----------|-------------------------------------|--------------------------------|------------------------------|----------------------------|
|          | 1<br>Preparation for<br>Instruction | 2<br>Use of Teaching<br>Skills | 3<br>Classroom<br>Management | 4<br>Personal<br>Qualities |
| 1        |                                     | 0.92**                         | 0.60**                       | 0.81**                     |
| 2        |                                     |                                | 0.72**                       | 0.84**                     |
| 3        |                                     |                                |                              | 0.67**                     |
| 4        |                                     |                                |                              |                            |

Sponsor Teachers

| Clusters | Clusters                            |                                |                              |                            |
|----------|-------------------------------------|--------------------------------|------------------------------|----------------------------|
|          | 1<br>Preparation for<br>Instruction | 2<br>Use of Teaching<br>Skills | 3<br>Classroom<br>Management | 4<br>Personal<br>Qualities |
| 1        |                                     | 0.67**                         | 0.60**                       | 0.65**                     |
| 2        |                                     |                                | 0.71**                       | 0.70**                     |
| 3        |                                     |                                |                              | 0.67**                     |
| 4        |                                     |                                |                              |                            |

\*  $p < .05$

\*\*  $p < .01$

To a lesser degree, student teachers also demonstrated global ratings in judging their own performance. As may be seen in Table 5, five of the six correlation coefficients for the student teachers achieved significance. However, the score for Cluster Three: Classroom Management was not significantly correlated with the student teachers' judgments of the remaining three clusters.

Possibly student teachers do tend to judge their performance in a relatively global manner, but single out the behavior of classroom management because it is of particular concern. This finding tends to support the findings in the Freeland (1979) and the Purcell and Seifert (1982) studies which noted that motivating disinterested pupils and handling discipline problems were student teachers' major concerns. The data from the present study do not indicate whether student teachers viewed classroom management as a major concern, but simply that they viewed their performance on this cluster differently from their performance on the other clusters of teaching behaviors. As a result of several informal conversations with students, it seems likely that classroom management was in fact a major concern. This speculation, combined with the results of the Freeland (1979) and Purcell and Seifert (1982) studies have definite implications for student teaching seminars. If classroom management is one of the student teacher's greatest concerns, both the university supervisor and the sponsor teacher must be sure to include ample information and instructions on this teaching behavior while conferring with students. Whether a generous amount of time is spent on the topic of classroom management is an issue for further research.

In the course of the study, a questionnaire of student teacher performance was constructed which may prove useful to other investigators, teacher trainers, and sponsor teachers. The instrument appears to be sufficiently reliable when used by university supervisors. However, further test development is necessary to improve the internal consistency of the questionnaire, when it is used by sponsor teachers or by student teachers to judge their own performance. In addition, although content validity was tentatively established through the methods reported above, estimates of empirical validity have yet to be measured. Following these steps, and the establishment of high inter-rater reliability coefficients, this questionnaire may be useful.

As this study was conducted at one institution and involved one student teacher training program, it is difficult to generalize the results to diverse populations of student teachers, university supervisors and sponsor teachers. Replication needs to be done in other institutions, in other training programs, and with student teachers who are not in their first practicum. As Borich and Madden (1977) have stated, the finding of no significant differences tends to be the norm in classroom research and more confidence can be placed in findings as they are replicated across studies using the same instrument. The study, however, has added to an extremely small body of research which has explored the question of whether members of the student teaching triad tend to judge the student teacher's performance as being at the same level of competence. In contrast to the Bueth and Pettibone (1973) study, this study used fairly inclusive measures of a student teacher's performance (i.e., Preparation for Instruction, Use of Teaching Skills, Classroom Management, and Personal Qualities); and in contrast to the Simmons (1973) study, this study used a larger sample (i.e., 50).

The study has also raised some interesting questions. How can university supervisors best communicate their expectations for lesson plans to sponsor teachers? Are most student teachers consistent in their degree of competence across

teaching behaviors or are they simply perceived by themselves and their supervisory personnel in such a manner? If in fact diverse competence exists, or if teaching behaviors are so interdependent that lack of competence in one behavior tends to radiate to lack of competence in other behaviors, how can supervisory personnel be trained to perceive specific problems and communicate them to student teachers? Do most student teachers perceive classroom management as their greatest weakness? If so, how much help is being provided by supervisory personnel in improving this teaching behavior? In general, it appears that student teachers, university supervisors, and sponsor teachers do tend to judge the student teacher's performance as being at the same level of competence, yet answers to the above questions would certainly help to increase the necessary communication within the student teaching triad.

### *References*

- Bartos, R., & Lotven, B. (1980). *Need and proficiency of teacher competencies as perceived by student teachers vs. practicing professional teachers*. (ERIC Document Reproduction Service No. ED 193 179)
- Borich, G., & Madden, S. (1977). *Evaluating classroom instruction: A sourcebook of instruments*. Reading, MA: Addison-Wesley.
- Buethel, C., & Pettibone, T. (1973). Evaluating student teaching behavior: Three views. *Teaching Education*, 9, 121-125.
- Campbell, L., & Williamson, J. (1973). Practical problems in the student teacher cooperative teacher relationship. *Education*, 94, 168-169.
- Chiarelott, L., Davidman, L., & Muse, C. (1980). Evaluating preservice teacher candidates. *The Clearing House*, 53, 295-299.
- Cooper, J. (Ed.). (1977). *Classroom teaching skills: A handbook*. New York: Harper & Row.
- Cronback, L. J. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.
- Freeland, K. (1979). Assessing student teaching. *Teacher Education*, 15, 11-16.
- Glass, G. V., & Stanley, J. C. (1970). *Statistical methods in education and psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Hattie, J., Olphert, W., & Cole, B. (1982). Assessment of student teachers by supervising teachers. *Journal of Educational Psychology*, 74, 778-785.
- Joyce, D. (1967). *Exploration of the utilization of personnel in the supervision of student teachers when feedback via films and systems for the analysis of teaching are introduced into the student teaching program* (Final Report). (ERIC Document Reproduction Service No. ED 020 575)
- Larson, E., & Ringness, T. (1965). *Effect of clinical evaluation on supervision of student teachers*. (ERIC Document Reproduction Service No. ED 003 467)
- Mahan, J., & Lester, F. (1974). *A study of inter-relationships between three independent rankings of student teacher effectiveness and student teacher assessments of satisfaction*. Chicago: Annual meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. ED 096 323)
- Nelson, L. (1974). *Guide to Letrap use and interpretation*. Dunedin, New Zealand: Department of Education, University of Otago.
- Peck, R., & Tucker, J. (1973). Research in teacher education. In R. Travers (Ed.), *The second handbook of research on teaching*. Chicago: Rand McNally.
- Purcell, T., & Seifert, B. (1982). A tri-state survey of student teachers. *College Student Journal*, 16, 27-29.



- Queen, J., & Mallen, L. (1982). Student teaching: A transitional approach. *The Clearing House*, 56, 56-59.
- Reynolds, J. (1977). A modified approach to teacher education assessment. *Teacher Education*, 13, 9-13.
- Rikard, G. (1982). The student teacher practicum. *Journal of Physical Education, Recreation and Dance*, 53, 60-61.
- Seiferth, B. M., & Solliday, M. A. (1980). A time study of the patterns of activities of student teachers. *College Student Journal*, 14, 285-287.
- Simmons, J. (1973). Two approaches: Name your poison. *English Education*, 4, 151-160.
- Tabachnik, B., Popkewitz, T., & Zeichner, K. (1980). Teacher education and the professional perspectives of student teachers. *Interchange*, 10, 12-19.
- Yee, A. (1969). Do cooperating teachers influence the attitudes of student teachers? *Journal of Educational Psychology*, 60, 328-332.

HORST H. MUELLER

*University of Alberta*

GABRIEL J. MANCINI

*Edmonton Public School Board*

and

ROBERT H. SHORT

*University of Alberta*

## An Evaluation of the Diagnostic Efficiency of the Wechsler Intelligence Scale for Children—Revised

*Although not intended as such, the WISC-R has been used extensively as a diagnostic instrument even though its effectiveness in such a role has remained in doubt. This study assessed the diagnostic suitability of the WISC-R by applying Kelley's (1923) method for judging the orthogonality of tests through the estimation of the proportion of score differences in excess of chance to all WISC-R subtest paired comparisons. Kelley's method was also applied to factor comparisons, Bannatyne (1974) cluster comparisons, and WISC-R Scale comparisons at each of the 11 age levels for which the WISC-R was standardized.*

*The results indicate that caution needs to be used when applying the WISC-R diagnostically. More than half of all possible subtest comparisons proved inappropriate for diagnostic purposes while others varied according to the age level in question. The factor, cluster, and Scale comparisons proved somewhat better for diagnosis. Generally, the findings of this study support the contention that three groupings of WISC-R subtests based on Kaufman's (1975) factor analysis are fully adequate for describing WISC-R performance. The authors advocate the use of the comparison tables presented in this paper as a necessary preliminary in any attempt to use the WISC-R as a diagnostic tool.*

---

Horst Mueller (MA, University of Saskatchewan) is presently a doctoral student in the Department of Educational Psychology at the University of Alberta. His research interests include behavior therapy with children, self-management training and behavioral medicine, clinical assessment, and reading comprehension processes.

Gabriel Mancini is a consulting psychologist with the Edmonton Public School Board, Edmonton, Alberta. His research interests are metacognition and behavioral disorder in children.

Robert Short is Associate Professor in the Department of Educational Psychology at the University of Alberta. His research interests include human learning and development, individual differences, and instructional psychology.

The Wechsler Intelligence Scale for Children-Revised (WISC-R) (Wechsler, 1974) is currently the most popular and well-researched psychometric instrument for the assessment of the intellectual functioning of individual children in both clinical and educational settings (Kaufman, 1980; Quattrocchi & Sherrets, 1980; Sattler, 1982). This popularity has come about in part because, although designed solely as a measure of "global intelligence" or an individual's "overall capacity for intelligent behavior" (Wechsler, 1974, p. 1), the WISC-R also has a number of features which render it attractive for use as a psychodiagnostic tool. Such features as the excellent reliability of its three Scale scores; the "adequate" reliable specific variance of most of its subtests (Dinning, 1976; Kaufman, 1980); the relative stability and remarkable similarity of its factor composition across age (Kaufman, 1975), sex, (Reynolds & Gutkin, 1980), intelligence levels (Karnes & Brown, 1980; Van Hagen & Kaufman, 1975), ethnic/racial groupings (Reschley, 1978), and clinical diagnostic groupings (Kaufman, 1979, 1981; Lutey, 1977; Quattrocchi & Sherrets, 1980; Sattler, 1982); together with its division into separate verbal and nonverbal scales have all greatly added to the WISC-R's clinical popularity.

In recent years, more and more research has focused on the WISC-R's potential as an educational and clinical differential diagnostic tool, and many clinicians have come to view the pattern (high versus low scores) of an individual's performance on the WISC-R subtests as potentially indicative of specific mental operations or brain-behavior relationships (Kaufman, 1980; Sattler, 1982). In this regard a considerable literature has accumulated on various WISC-R test scatter and subtest profile indices and their association with specific exceptionalities (e.g., Anderson, Kaufman, & Kaufman, 1976; Bannatyne, 1968, 1974; Clarizio & Bernard, 1981; Dash, Dennis, Mueller, Mancini, Snart, & Short, 1983; Dean, 1977, 1978; Groff & Hubble, 1981; Hale, 1979; Hale & Landino, 1981; Haynes & Bensch, 1981; Kuncie & McMahon, 1979; Loro & Woodward, 1976; McIntosh, 1974; Mueller, Matheson, & Short, 1983; Tew, 1977; Thompson, 1980; Thompson, 1981; Vance & Singer, 1979; Webster & Lafayette, 1980). Unfortunately, the results of such studies have been mixed and have failed to provide any firm basis for making diagnostic classification decisions on the basis of WISC-R subtest scatter or performance profiles (Kaufman, 1981; Lutey, 1977; Sattler, 1982). A number of authors have also leveled criticism at both the usefulness of such research and at some of the basic assumptions underlying the use of scatter and profile analysis approaches to differential diagnosis and classification of children with the WISC-R (e.g., Gutkin, 1979a; Hirshoren & Kavale, 1976; Miller, 1980; Ryckman, 1981). A major caution stated by many authors is that the reliabilities of the individual WISC-R subtests are too low to warrant interpretation of performance profiles based on individual subtest scores (Hirshoren & Kavale, 1976). However, the practice of WISC-R profile analysis, based on either individual subtests or clusters of subtests, continues among both researchers and clinicians (Kaufman, 1979; Sattler, 1982).

The problem of diagnosis may be seen as twofold. Diagnosis is both the process of classifying individuals on the basis of their problems or difficulties, and the process of determining the nature of the problem or difficulty so as to guide appropriate intervention or remediation (see Chaplin, 1975). Similarly, a diagnostic test is any instrument which is utilized in an effort to classify individuals on the basis of the difficulties they demonstrate and/or to determine the nature or source of such individuals' problems (Salvia & Ysseldyke, 1981). The WISC-R has been used both in attempts to classify individuals on the basis of their subtest performance profiles (e.g., Vance, Wallbrown, & Blaha, 1978; Vance & Singer, 1979), and to determine



specific areas of difficulty and suggest remediation strategy for children already classified (e.g., Wallbrown, Vance, & Blaha, 1979). The first of these usages involves inter-individual comparisons of subtest performance profiles, whereas the second usage involves intra-individual comparisons. Thus the WISC-R must be capable of determining "true" differences in performance profiles of individuals from the norm and, at the same time, it must be able to determine "true" or reliable intra-individual differences in abilities. In a sense, any normative-based psychodiagnostic instrument must be adequate for both these uses.

A major failing within this research area has been the lack of studies examining whether or not the WISC-R, which was designed to reliably differentiate individuals on the basis of its three Scale scores, is indeed capable of differentiating individuals on the basis of differential scores across subtest pairs or across clusters of subtests based on rational (Bannatyne, 1968, 1974) or factorial (Kaufman, 1975) grouping schema. This is an important question because, underlying the profile analysis research is the often less-than-explicit notion that a specific pattern of differential WISC-R performance that is reliable for the individual will also reliably distinguish that individual from the normative group. The usual approach to using the WISC-R diagnostically has been to simply examine the individual's profile of subtest or subtest-cluster scores for score differences large enough to be significant at the .05 or .01 level of probability as determined from various published tables of significant score discrepancies (Gutkin, 1978, 1979b; Kaufman, 1979; Lutey, 1977; Piotrowski & Grubb, 1976; Reynolds, 1981; Sattler, 1982). The application of such procedures to the test results of an individual assumes that the subtests or subtest-clusters have been found, by an analysis of results for the population, capable of reliably showing such differences. Such an analysis has yet to be done for the WISC-R.

For any test or test-battery to demonstrate adequate diagnostic efficiency it must be comprised of subtests which are not highly intercorrelated and which are individually quite reliable. Such a combination of psychometric properties will ensure that the test is capable of measuring relatively independent traits and that any obtained score differences between subtests will be attributable to true individual differences in the traits being measured rather than to measurement error (Salvia & Ysseldyke, 1981). The WISC-R seems a rather dubious candidate for an efficient diagnostic test because it is composed of subtests that have a relatively large amount of common variance with less individual specific variance (Dinning, 1976; Sattler, 1982), and its individual subtests vary greatly in reliability (Wechsler, 1974).

The present paper is concerned specifically with an analysis and evaluation of the differential diagnostic efficiency of the WISC-R and will apply Kelley's (1923) method for estimating the proportion of differences in excess of chance to all possible WISC-R subtest paired comparisons, Scale comparisons, Kaufman (1975) factor comparisons, and Bannatyne (1974) cluster comparisons at each of the 11 age levels for which the WISC-R was standardized.

Although other procedures for assessing the reliability of score differences between tests are available (Mollenkopf, 1950; Mosier, 1951; Nunnally, 1967), Kelley's method for determining the proportion of differences in excess of chance has the advantage of being directly meaningful and easily understandable (French, 1966). Unfortunately, apart from its use in studies by Segel (1934), Bennett and Doppelt (1948), and Tierney and Ames (1978), this particular technique has been rarely ap-

plied to test data. Based on the reliabilities of the tests and their intercorrelations, the index derived by Kelley's method is equivalent to the percentage of time differences between two tests represent true differences in ability.

A better understanding of Kelley's approach for estimating the diagnostic efficiency of two tests or batteries of tests may be gained by supposing that two persons stand at exactly the same level on some aptitude. When these two people are tested for this aptitude by fallible tests, there will be a difference in the scores they receive. If such testing is done repeatedly, a distribution of difference scores will be obtained. This distribution, of course, will be entirely attributable to chance. In cases where real differences in aptitude exist between two people, the observed differences in scores will be greater and will be partly attributable to chance and partly attributable to real differences in aptitude. The efficiency of a diagnostic test can be stated in terms of the proportion of observed differences that are not attributable to chance. If the two variables in question are highly related, the real differences will be rather small and, therefore, the proportion of differences not accounted for by chance will be low. If, on the other hand, the two variables are highly independent, the differences will be larger. Test reliability also comes into play here in that if the two tests have poor reliability, a greater proportion of differences will be due to chance. Thus, as stated earlier, a good differential diagnostic test must be comprised of subtests that are highly independent of each other and highly reliable. Kelley's (1923) formula permits a determination of the diagnostic efficiency of a test through the determination of that proportion of differences between two subtests or batteries of tests, based on the tests' intercorrelation and reliabilities, that is in excess of chance. For example, for two tests a proportion of differences in excess of chance of .35 indicates that fully 65 percent of all score differences would be due to measurement error. The larger the proportion of differences in excess of chance, the more diagnostically efficient the test.

Segel (1934), in his research with Kelley's procedure, concluded that tests exhibiting a proportion of score differences in excess of chance less than .25 should not be considered diagnostic in relation to each other—such tests having little value for differential diagnosis. Segel's minimum criterion for diagnostic efficiency may at first glance appear overly lenient, since at the .25 level he advocates fully 75 percent of all score differences between two tests would be due to chance. However, a review of the few studies which have examined the diagnostic efficiency of tests using the Kelley method (Bennett & Doppelt, 1948; Bennett, Seashore, & Wesman, 1947; Kelley, 1923; Segel, 1934; Tierney & Ames, 1978) reveals that few tests are capable of differentiating abilities at much better than 25 percent above chance. For example, Kelley (1923) discovered that the Stanford Achievement Test was capable of differentiating arithmetic (total arithmetic score) from reading (total reading score) abilities at no more than 36 percent above chance. Similarly, Segel (1934) found that comparisons between the three parts of the Gates Primary Reading Test and the four parts of the Gates Silent Reading Test demonstrated a mean diagnostic efficiency of barely more than 25 percent above chance. More recently, Tierney and Ames (1978), using the Kelley procedure to evaluate the diagnostic efficiency of the Illinois Test of Psycholinguistic Abilities, reported "that 45 out of the 66 possible subtest comparisons were . . . inappropriate [below .25] at at least one age level" (p. 589). Such research results clearly show the Segel criterion for diagnostic efficiency to be not overly liberal.



### *Method*

The Kelley (1923) procedure for estimating the proportion of score differences between two tests in excess of chance was applied to the WISC-R standardization data (Wechsler, 1974). Using Segel's (1934) criterion of 25 percent of score differences in excess of chance as a minimum standard for diagnostic utility, the diagnostic efficiency of WISC-R subtest paired comparisons, WISC-R Scale comparisons, Kaufman (1975) factor grouping comparisons, and Bannatyne (1974) cluster comparisons were evaluated for each of the 11 age groups of the standardization population.

First, using the subtest reliabilities and intercorrelation coefficients given in the WISC-R manual (pp. 28, 36-47), Kelley's formula for estimating tests' diagnostic efficiency was applied to each of the 66 unique WISC-R subtest paired comparisons at each of 11 age levels and the proportion of differences in excess of chance was calculated for each of these 726 comparisons. A similar set of calculations was carried out for the WISC-R Verbal Scale versus Performance Scale comparisons at each of the 11 age levels using the Scale reliabilities and intercorrelation coefficients published in the WISC-R manual.

For the evaluation of the diagnostic efficiency of the three WISC-R factor groupings of subtests first derived by Kaufman (1975), a more complex procedure was followed. First, a composite reliability was estimated for each of the three factor groupings of subtests at each of 11 age levels using the general formula developed by Mosier (Guilford, 1954, p. 393). The intercorrelation of the three factor groupings of subtests at all 11 age levels was also estimated by transforming the composite subtest intercorrelation coefficients given in the WISC-R manual to Fisher Z-scores (Smith, 1970), obtaining their mean for each factor grouping, and then converting this mean Z-score back into an intercorrelation coefficient. These composite reliabilities and intercorrelation estimates for the three factor groupings of subtests were then used in Kelley's formula to calculate the proportion of score differences in excess of chance for all possible paired comparisons of the Verbal Comprehension (Information, Similarities, Vocabulary, Comprehension), Perceptual Organization (Picture Completion, Picture Arrangement, Block Design, Object Assembly), and Freedom from Distractibility (Arithmetic, Digit Span, Coding) factors at each of 11 age levels.

Lastly, a similar procedure to that used to estimate the diagnostic efficiency of Kaufman's WISC-R factors was used to estimate the diagnostic efficiency of the four Bannatyne (1974) WISC-R subtest clusters. As before, the proportion of score differences in excess of chance was calculated for all possible paired comparisons of the Bannatyne clusters at each of 11 age levels.

### *Results*

Table 1 presents the proportion, expressed as a percentage, of score differences in excess of chance for WISC-R subtest paired-comparisons at each of the 11 age levels for which the WISC-R was standardized. From this table it can be seen that each of the individual WISC-R subtests are involved in at least 38 inappropriate comparisons. At certain age levels, some of the subtests (e.g., Information at the 6.5 year level) are never involved in an appropriate comparison. The subtests which fare best are Digit Span, Coding, and Block Design where 69 percent, 58 percent, and 50 percent of comparisons, respectively, meet or exceed Segel's .25 criterion for minimal diagnostic efficiency. On the other hand, Object Assembly, Comprehen-



TABLE 1  
PERCENTAGE OF SCORE DIFFERENCES IN EXCESS OF CHANCE  
FOR ALL WISC-R SUBTEST PAIRED-COMPARISONS

| Subtest Pairs | Age Level |     |     |     |      |      |      |      |      |      |      |
|---------------|-----------|-----|-----|-----|------|------|------|------|------|------|------|
|               | 6.5       | 7.5 | 8.5 | 9.5 | 10.5 | 11.5 | 12.5 | 13.5 | 14.5 | 15.5 | 16.5 |
| Info.-Simil.  | 21        | 22  | 17  | 12  | 18   | 16   | 21   | 16   | 17   | 17   | 24   |
| Info.-Arith.  | 11        | 13  | 19  | 19  | 21   | 24   | 26*  | 24   | 16   | 30*  | 21   |
| Info.-Voc.    | 11        | 14  | 21  | 12  | 19   | 17   | 19   | 17   | 19   | 23   | 22   |
| Info.-Comp.   | 15        | 17  | 18  | 18  | 18   | 23   | 26*  | 21   | 23   | 22   | 22   |
| Info.-D.Sp.   | 15        | 27* | 34* | 22  | 26*  | 29*  | 29*  | 31*  | 31*  | 35*  | 30*  |
| Info.-P.Com.  | 23        | 28* | 32* | 26* | 23   | 29*  | 23   | 28*  | 23   | 25*  | 27*  |
| Info.-P.Arr.  | 14        | 19  | 18  | 23  | 26*  | 25*  | 25*  | 23   | 25*  | 32*  | 28*  |
| Info.-B.Des.  | 17        | 23  | 29* | 21  | 31*  | 32*  | 28*  | 29*  | 30*  | 36*  | 37*  |
| Info.-O.Ass.  | 19        | 18  | 21  | 21  | 19   | 24   | 19   | 30*  | 24   | 28*  | 25*  |
| Info.-Cod.    | 17        | 23  | 21  | 28* | 32*  | 31*  | 34*  | 24   | 27*  | 37*  | 32*  |
| Info.-Maz.    | 21        | 29* | 31* | 27* | 26*  | 30*  | 22   | 32*  | 31*  | 28*  | 24   |
| Simil.-Arith. | 29*       | 25* | 21  | 22  | 19   | 22   | 26*  | 18   | 17   | 21   | 23   |
| Simil.-Voc.   | 23        | 14  | 21  | 10  | 11   | 14   | 19   | 12   | 15   | 14   | 22   |
| Simil.-Comp.  | 23        | 21  | 15  | 13  | 13   | 17   | 21   | 14   | 13   | 11   | 17   |
| Simil.-D.Sp.  | 27*       | 33* | 32* | 18  | 24   | 30*  | 28*  | 27*  | 25*  | 24   | 29*  |
| Simil.-P.Com. | 33*       | 31* | 30* | 22  | 17   | 24   | 21   | 20   | 17   | 15   | 20   |
| Simil.-P.Arr. | 28*       | 23  | 21  | 18  | 22   | 24   | 26*  | 20   | 20   | 19   | 25*  |
| Simil.-B.Des. | 32*       | 27* | 27* | 16  | 24   | 26*  | 26*  | 25*  | 23   | 23   | 31*  |
| Simil.-O.Ass. | 30*       | 23  | 21  | 14  | 15   | 19   | 17   | 24   | 19   | 19   | 24   |
| Simil.-Cod.   | 29*       | 27* | 14  | 27* | 26*  | 27*  | 32*  | 21   | 23   | 26*  | 28*  |
| Simil.-Maz.   | 33*       | 35* | 24  | 23  | 24   | 26*  | 21   | 27*  | 25*  | 19   | 21   |
| Arith.-Voc.   | 17        | 14  | 25* | 21  | 23   | 26*  | 24   | 25*  | 22   | 28*  | 25*  |
| Arith.-Comp.  | 18        | 17  | 21  | 22  | 21   | 25*  | 27*  | 22   | 20   | 21   | 21   |
| Arith.-D.Sp.  | 19        | 23  | 23  | 19  | 21   | 24   | 19   | 24   | 21   | 25*  | 17   |
| Arith.-P.Com. | 29*       | 29* | 27* | 27* | 22   | 27*  | 24   | 22   | 21   | 22   | 25*  |
| Arith.-P.Arr. | 23        | 21  | 19  | 27* | 24   | 26*  | 27*  | 27*  | 22   | 28*  | 25*  |
| Arith.-B.Des. | 23        | 24  | 24  | 21  | 28*  | 28*  | 22   | 27*  | 23   | 28*  | 30*  |
| Arith.-O.Ass. | 23        | 20  | 19  | 24  | 21   | 24   | 19   | 28*  | 23   | 27*  | 26*  |
| Arith.-Cod.   | 21        | 22  | 14  | 25* | 25*  | 30*  | 27*  | 22   | 21   | 30*  | 25*  |
| Arith.-Maz.   | 28*       | 28* | 24  | 25* | 21   | 25*  | 18   | 26*  | 26*  | 24   | 19   |
| Voc.-Comp.    | 12        | 06  | 15  | 14  | 13   | 16   | 24   | 16   | 17   | 14   | 19   |
| Voc.-D.Sp.    | 23        | 24  | 40* | 22  | 28*  | 29*  | 26*  | 30*  | 30*  | 30*  | 30*  |
| Voc. P.Com.   | 23        | 23  | 31* | 28* | 19   | 27*  | 24   | 28*  | 26*  | 25*  | 27*  |
| Voc.-P.Arr.   | 18        | 17  | 20  | 23  | 24   | 26*  | 24   | 25*  | 26*  | 30*  | 28*  |
| Voc.-B.Des.   | 21        | 21  | 34* | 24  | 28*  | 32*  | 25*  | 31*  | 33*  | 36*  | 39*  |
| Voc.-O.Ass.   | 21        | 20  | 25* | 20  | 17   | 24   | 18   | 29*  | 28*  | 30*  | 28*  |
| Voc.-Cod.     | 20        | 18  | 24  | 28* | 30*  | 32*  | 32*  | 24   | 28*  | 34*  | 34*  |
| Voc.-Maz.     | 27*       | 26* | 33* | 28* | 27*  | 31*  | 21   | 32*  | 32*  | 28*  | 24*  |

sion, and Similarities fare least well with only 20 percent, 28 percent, and 31 percent of comparisons being diagnostically tenable. A total of 11 of the 66 possible subtest paired-comparisons are diagnostically valueless at all age levels. These are:

TABLE 1 (CONT'D)

| Subtest Pairs | Age Level |     |     |     |      |      |      |      |      |      |      |
|---------------|-----------|-----|-----|-----|------|------|------|------|------|------|------|
|               | 6.5       | 7.5 | 8.5 | 9.5 | 10.5 | 11.5 | 12.5 | 13.5 | 14.5 | 15.5 | 16.5 |
| Comp.-D.Sp.   | 21        | 25* | 33* | 22  | 26*  | 29*  | 31*  | 28*  | 29*  | 30*  | 29*  |
| Comp.-P.Com.  | 24        | 23  | 27* | 19  | 17   | 26*  | 23*  | 19   | 16   | 16   | 21   |
| Comp.-P.Arr.  | 18        | 17  | 13  | 19  | 21   | 24   | 28*  | 23   | 21   | 21   | 24   |
| Comp.-B.Des.  | 19        | 21  | 27* | 20  | 25*  | 32*  | 31*  | 28*  | 27*  | 24   | 30*  |
| Comp.-O.Ass.  | 21        | 17  | 18  | 15  | 18   | 23   | 23   | 25*  | 23   | 20   | 22   |
| Comp.-Cod.    | 21        | 21  | 19  | 28* | 26*  | 30*  | 32*  | 23   | 24   | 26*  | 26*  |
| Comp.-Maz.    | 21        | 23  | 23  | 26* | 21   | 28*  | 24   | 27*  | 28*  | 18   | 17   |
| D.Sp.-P.Comp. | 30*       | 36* | 34* | 25* | 27*  | 32*  | 24   | 32*  | 28*  | 27*  | 29*  |
| D.Sp.-P.Arr.  | 23        | 26* | 32* | 27* | 25*  | 27*  | 30*  | 27*  | 24   | 29*  | 29*  |
| D.Sp.-B.Des.  | 25*       | 32* | 37* | 24  | 26*  | 34*  | 26*  | 32*  | 33*  | 32*  | 35*  |
| D.Sp.-O.Ass.  | 25*       | 28* | 28* | 23  | 21   | 27*  | 22   | 31*  | 26*  | 27*  | 28*  |
| D.Sp.-Cod.    | 21        | 27* | 22  | 23  | 22   | 24   | 27*  | 19   | 24   | 30*  | 30*  |
| D.Sp.-Maz.    | 28*       | 32* | 34* | 25* | 22   | 31*  | 21   | 25*  | 27*  | 24   | 22   |
| P.Com.-P.Arr. | 23        | 24  | 24  | 19  | 21   | 23   | 16   | 21   | 21   | 17   | 22   |
| P.Com.-B.Des. | 24        | 27* | 32* | 18  | 19   | 27*  | 18   | 18   | 16   | 17   | 22   |
| P.Com.-O.Ass. | 24        | 21  | 18  | 14  | 10   | 16   | 12   | 18   | 14   | 16   | 15   |
| P.Com.-Cod.   | 23        | 27* | 23  | 29* | 28*  | 35*  | 30*  | 24   | 23   | 28*  | 30*  |
| P.Com.-Maz.   | 32*       | 26* | 31* | 21  | 17   | 27*  | 15   | 21   | 19   | 16   | 18   |
| P.Arr.-B.Des. | 18        | 18  | 21  | 21  | 26*  | 28*  | 25*  | 21   | 22   | 24   | 26*  |
| P.Arr.-O.Ass. | 18        | 14  | 10  | 16  | 16   | 18   | 14   | 19   | 21   | 19   | 18   |
| P.Arr.-Cod.   | 19        | 18  | 17  | 27* | 26*  | 26*  | 29*  | 23   | 21   | 30*  | 28*  |
| P.Arr.-Maz.   | 24        | 24  | 21  | 23  | 20   | 24   | 17   | 23   | 21   | 19   | 17   |
| B.Des.-O.Ass. | 18        | 11  | 13  | 10  | 12   | 13   | 13   | 16   | 13   | 13   | 14   |
| B.Des.-Cod.   | 19        | 23  | 21  | 28* | 30*  | 33*  | 29*  | 23   | 26*  | 30*  | 34*  |
| B.Des.-Maz.   | 24        | 23  | 28* | 18  | 19   | 27*  | 20   | 21   | 24   | 19   | 16   |
| O.Ass.-Cod.   | 19        | 23  | 17  | 24  | 24   | 28*  | 23   | 21   | 22   | 26*  | 26*  |
| O.Ass.-Maz.   | 18        | 21  | 17  | 17  | 19   | 24   | 11   | 17   | 22   | 16   | 14   |
| Cod.-Maz.     | 27*       | 23  | 23  | 26* | 24   | 32*  | 24   | 19   | 19   | 25*  | 22   |

Note:Asterisks denote those comparisons with a diagnostic efficiency of .25 or better.

Information with either Similarities or Vocabulary, Similarities with either Vocabulary or Comprehension, Vocabulary with Comprehension, Picture Completion with either Picture Arrangement or Object Assembly, Picture Arrangement with either Object Assembly or Mazes, and Object Assembly with either Block Design or Mazes. Generally, score differences between individual subtest-pairs within either the Verbal Comprehension factor or the Perceptual Organization factor (Kaufman, 1975) are most likely to be due to chance. All of the 66 possible subtest paired-comparisons failed to meet Segel's minimal requirement for diagnostic efficiency at at least one age level. Clearly, basing diagnostic statements upon score differences between individual WISC-R subtests is a rather risky proposition.

TABLE 2  
PERCENTAGE OF SCORE DIFFERENCES IN EXCESS OF CHANCE FOR  
WISC-R SCALE COMPARISONS, KAUFMAN FACTOR  
COMPARISONS, AND BANNATYNE CLUSTER COMPARAISONS

| WISC-R Scales      | Age Level |     |     |     |      |      |      |      |      |      |      |
|--------------------|-----------|-----|-----|-----|------|------|------|------|------|------|------|
|                    | 6.5       | 7.5 | 8.5 | 9.5 | 10.5 | 11.5 | 12.5 | 13.5 | 14.5 | 15.5 | 16.5 |
| Verb.-Perf.        | 32*       | 32* | 33* | 31* | 31*  | 34*  | 33*  | 36*  | 32*  | 37*  | 36*  |
| Kaufman Factors    |           |     |     |     |      |      |      |      |      |      |      |
| VC-PO              | 39*       | 43* | 41* | 37* | 41*  | 45*  | 45*  | 45*  | 44*  | 41*  | 47*  |
| VC-FD              | 41*       | 42* | 47* | 45* | 41*  | 45*  | 51*  | 45*  | 45*  | 47*  | 53*  |
| PO-FD              | 42*       | 45* | 42* | 45* | 36*  | 47*  | 47*  | 44*  | 45*  | 45*  | 51*  |
| Bannatyne Clusters |           |     |     |     |      |      |      |      |      |      |      |
| Con.-Spat.         | 40*       | 37* | 42* | 37* | 36*  | 43*  | 39*  | 42*  | 42*  | 39*  | 41*  |
| Con.-Seq.          | 36*       | 36* | 40* | 39* | 39*  | 42*  | 42*  | 39*  | 37*  | 41*  | 42*  |
| Con.-Ac.Kn.        | 16        | 08  | 14  | 12  | 13   | 08   | 13   | 14   | 18   | 16   | 16   |
| Spat.-Seq.         | 37*       | 39* | 37* | 39* | 38*  | 43*  | 39*  | 40*  | 37*  | 41*  | 43*  |
| Spat.-Ac.Kn.       | 37*       | 37* | 40* | 41* | 39*  | 43*  | 39*  | 44*  | 42*  | 44*  | 43*  |
| Seq.-Ac.Kn.        | 09        | 08  | 15  | 14  | 16   | 19   | 14   | 17   | 13   | 22   | 16   |

Note:

1. Asterisks denote comparisons with a diagnostic efficiency of .25 or better.
2. VC = Verbal Comprehension, PO = Perceptual Organization, FD = Freedom from Distractibility, Con. = Conceptual, Spat. = Spatial, Seq = Sequential, and Ac.Kn. = Acquired Knowledge

An examination of the diagnostic efficiency of various groupings or clusterings of WISC-R subtests presents a somewhat brighter outlook than that for individual subtest pairings. Table 2 presents the percentage of score differences in excess of chance for WISC-R Scale comparisons, factor comparisons, and Bannatyne cluster comparisons at each of 11 age levels. From this table it can be seen that comparisons between the WISC-R Verbal Scale and Performance Scale are diagnostically tenable at all age levels with an efficiency of at least 30 percent above chance. This apparent diagnostic efficiency of the WISC-R's two Scales is even further improved when the WISC-R is divided into three scales based on Kaufman's (1975) factor analytic research with the WISC-R. The use of a three-scale interpretation permits the differentiation of abilities at better than 40 percent above chance for most of the possible comparisons.

An examination of Table 2 also reveals that for the four Bannatyne clusters of subtests only certain paired comparisons are tenable at all age levels, whereas others are not. The three Bannatyne clusters that most closely resemble the Kaufman WISC-R factor groupings, that is, the Conceptual, Spatial, and Sequential clusters, all demonstrate adequate diagnostic efficiency at all age levels when paired with each other. On the other hand, the fourth Bannatyne cluster, Acquired Knowledge, lacks diagnostic value when involved in comparisons with either



the Conceptual or Sequential clusters. This lack of diagnostic efficiency for such pairings is a function of the fact that Acquired Knowledge contains subtests that are also found in the Conceptual and Sequential clusters, thus dramatically increasing the intercorrelation of Acquired Knowledge with both Conceptual and Sequential (Matheson, Mueller, & Short, in press). Table 2 also reveals that, generally, paired comparisons between Bannatyne clusters of subtests are less diagnostically efficient than paired comparisons between Kaufman factor groupings at most age levels. This is due to the somewhat lower composite reliability of the Bannatyne clusters.

### *Discussion*

The analysis of WISC-R standardization data (Wechsler, 1974) using Kelley's (1923) procedure for the determination of the proportion of all score differences between tests in excess of chance permitted an evaluation of the diagnostic efficiency of the WISC-R subtests and Scales, Kaufman's (1975) WISC-R factor groupings, and Bannatyne's (1974) recategorized clusters of WISC-R subtests. The application of Segel's (1934) criterion for minimal diagnostic efficiency to the results of the data generated by Kelley's procedure served to reference those subtest, factor group, and cluster comparisons that may be considered diagnostically tenable at the age levels for which the WISC-R was originally standardized. Assuming the logic of our analyses and procedures, reference to Tables 1 and 2 should precede any attempt to use the WISC-R in a diagnostic manner. If, according to the data presented in our tables, certain subtest or cluster comparisons are diagnostically valuable, then appropriate standard error of measurement confidence intervals (see Gutkin, 1978; Reynolds, 1981; Sattler, 1982) may be applied to determine if a score difference is statistically significant. However, while the acceptance of the value of a particular subtest or cluster comparison does not negate taking standard differential diagnostic precautions (see Sattler, 1982), the rejection of a particular comparison on grounds of poor diagnostic efficiency does suggest that the compared tests cannot accurately discriminate real inter-individual differences in measured abilities from score differences caused by measurement error, regardless of the size of such differences. The chances of statistically significant differences being caused by measurement error are far too great for test comparisons with a diagnostic efficiency of less than 25 percent to have any value in differential diagnosis. For these reasons, reference to the tables presented herein should be a first step in any attempt to use the WISC-R as a diagnostic tool.

More generally, the results of the present analysis lend further support to previously expressed cautions against differential diagnosis based on WISC-R subtest comparisons (e.g., Hirshoren & Kavale, 1976; Lutey, 1977; Sattler, 1982; Wechsler, 1974), and to those who argue that three groupings of subtests are fully adequate for describing performance on the WISC-R (Hirshoren & Kavale, 1976; Kaufman, 1975). While the Bannatyne clusters appear to permit four tenable comparisons compared to only three using the Kaufman factor groupings, the Kaufman groupings show the highest diagnostic efficiency at all age levels. Certainly the four-cluster Bannatyne (1974) recategorization of the WISC-R subtests would appear to add little to the Kaufman three-factor solution other than permitting a tenable comparison between the Spatial and Acquired Knowledge clusters. This finding was given further confirmation recently when multiple group factor analysis was used to test Bannatyne's (1974) categories against the subtest correlations for the 7.5, 10.5,

and 13.5 year samples from the WISC-R standardization. In Matheson, Mueller, and Short's (in press) study, it was concluded "that Bannatyne's Acquired Knowledge category has no validity as a Construct which is separate from his conceptualization category" (p. 22).

In conclusion, it seems clear that although the WISC-R was not designed as a diagnostic test, it does show some efficiency in differentiating children on the basis of at least three broad abilities or aptitudes. However, in most cases, diagnostic attempts based on score differences between individual WISC-R subtests must be viewed as extremely poor practice.

#### Notes

1. Reprints may be obtained from the first author at the Department of Educational Psychology, Faculty of Education, University of Alberta, Edmonton, Alberta, Canada.
2. This paper was partially funded by a Social Sciences and Humanities Research Council of Canada Doctoral Fellowship awarded to the first author.
3. The authors wish to acknowledge the assistance of Udaya Dash, Steven Dennis, and Dwight Harley for their help in data analysis and programing, and editing.

#### References

- Anderson, M., Kaufman, A., & Kaufman, N. (1976). Use of the WISC-R with a learning disabled population: Some diagnostic implications. *Psychology in the Schools*, 13(3), 381-386.
- Bannatyne, A. (1968). Diagnosing learning disabilities and writing remedial prescriptions. *Journal of Learning Disabilities*, 1(4), 28-35.
- Bannatyne, A. (1974). Diagnosis: A note on recategorization of the WISC-R scaled scores. *Journal of Learning Disabilities*, 7(2), 272-274.
- Bennett, G., & Doppelt, J. (1948). The evaluation of pairs of tests for guidance use. *Educational & Psychological Measurement*, 8(3), 319-325.
- Bennett, G., Seashore, H., & Wesman, J. (1947). *Differential aptitude tests*. New York, NY: Psychological Corporation.
- Chaplin, J. (1975). *Dictionary of psychology* (revised ed.). New York, NY: Dell.
- Clarizio, H., & Bernard, R. (1981). Recategorized WISC-R scores of LD children and differential diagnosis. *Psychology in the Schools*, 18(1), 5-12.
- Dash, U., Dennis, S., Mueller, H., Mancini, G., Snart, F., & Short, R. (1983). WISC-R subtest variability in a sample of clinic-referred Canadian children. *Canadian Journal of Behavioral Science*, 15(3), 211-227.
- Dean, R. (1977). Patterns of emotional disturbance on the WISC-R. *Journal of Clinical Psychology*, 33(3), 486-490.
- Dean, R. (1978). Distinguishing learning disabled and emotionally disturbed children on the WISC-R. *Journal of Consulting & Clinical Psychology*, 46(2), 381-382.
- Dinning, W. (1976). Variance components of the WISC-R in eleven age groups. *Psychological Reports*, 38, 1001-1002.
- French, J. (1966). The logic of and assumptions underlying differential testing. In A. Anastasi (Ed.), *Testing problems in perspective*. Washington, DC: American Council on Education.
- Groff, M., & Hubble, L. (1981). Recategorized WISC-R scores of juvenile delinquents. *Journal of Learning Disabilities*, 14(7), 519-520.
- Guilford, J. (1954). *Psychometric methods* (2nd ed.). New York, NY: McGraw-Hill.
- Gutkin, T. (1978). Some useful statistics for the interpretation of the WISC-R. *Journal of Consulting & Clinical Psychology*, 46(6), 1561-1563.



- Gutkin, T. (1979a). WISC-R scatter indices: Useful information for differential diagnoses? *Journal of School Psychology, 17*(4), 368-371.
- Gutkin, T. (1979b). The WISC-R Verbal Comprehension, Perceptual Organization, and Freedom from Distractibility deviation quotients: Data for practitioners. *Psychology in the Schools, 16*(3), 359-360.
- Hale, R. (1979). The utility of WISC-R subtest scores in discriminating among adequate and underachieving children. *Multivariate Behavioral Research, 14*(2), 245-253.
- Hale, R., & Landino, S. (1981). Utility of WISC-R subtest analysis in discriminating among groups of conduct problem, withdrawn, mixed, and nonproblem boys. *Journal of Consulting & Clinical Psychology, 49*(1), 91-95.
- Haynes, J., & Bensch, M. (1981). The P > V sign on the WISC-R and recidivism in delinquents. *Journal of Consulting & Clinical Psychology, 49*(3), 480-481.
- Hirshoren, A., & Kavale, K. (1976). Profile analysis of the WISC-R: A continuing malpractice. *Exceptional Children, 23*(1), 83-87.
- Karnes, F., & Brown, K. (1980). Factor analysis of the WISC-R for the gifted. *Journal of Educational Psychology, 72*(2), 197-199.
- Kaufman, A. (1975). Factor analysis of the WISC-R. *Journal of Consulting & Clinical Psychology, 43*(1), 135-147.
- Kaufman, A. (1979). *Intelligent testing with the WISC-R*. New York, NY: J. Wiley & Sons.
- Kaufman, A. (1980). Issues in psychological assessment: Interpreting the WISC-R intelligently. In B. Lahey & A. Kazdin (Eds.), *Advances in clinical child psychology*. New York, NY: Plenum.
- Kaufman, A. (1981). The WISC-R and learning disabilities assessment: State of the art. *Journal of Learning Disabilities, 14*(9), 520-526.
- Kelley, T. (1923). A new method for determining the significance of differences in intelligence and achievement scores. *Journal of Educational Psychology, 14*(6), 321-333.
- Kunce, J., & McMahon, R. (1979). Neuropsychological significance of children's Wechsler intelligence scores. *Psychological Reports, 44*(3, pt. 1), 787-790.
- Loro, B., & Woodward, J. (1976). Verbal and performance IQ for discrimination among psychiatric diagnostic groups. *Journal of Clinical Psychology, 32*(1), 107-114.
- Lutey, C. (1977). *Individual intelligence testing: A sourcebook*. Greeley, CO: Lutey.
- Matheson, D., Mueller, H., & Short, R. (in press). The validity of Bannatyne's Acquired Knowledge category as a separate construct. *Journal of Psychoeducational Assessment*.
- McIntosh, W. (1974). The use of a Wechsler subtest ratio as an index of brain damage in children. *Journal of Learning Disabilities, 7*(2), 161-163.
- Miller, M. (1980). On the attempt to find WISC-R profiles for learning and reading disabilities. *Journal of Learning Disabilities, 13*(6), 52-54.
- Mollenkopf, W. (1950). Predicted differences and differences between predictions. *Psychometrika, 15*(4), 409-417.
- Mosier, C. (1951). Batteries and profiles. In E. F. Lindquist (Ed.), *Educational measurement*. Washington, DC: American Council on Education.
- Mueller, H., Matheson, D., & Short, R. (1983). Bannatyne-recategorized WISC-R patterns of mentally retarded, learning disabled, normal, and intellectually superior children: A meta-analysis. *The Mental Retardation and Learning Disability Bulletin, 11*(2), 60-78.
- Nunnally, J. (1967). *Psychometric theory*. New York, NY: McGraw-Hill.
- Piotrowski, R., & Grubb, R. (1976). Significant subtest score differences on the WISC-R. *Journal of School Psychology, 14*(2), 202-206.
- Quattrocchi, M., & Sherrets, S. (1980). WISC-R: The first five years. *Psychology in the Schools, 17*(3), 297-312.
- Reschley, D. (1978). WISC-R factor structures among Anglos, Blacks, Chicanos, and native American Papagos. *Journal of Consulting & Clinical Psychology, 46*, 417.



- Reynolds, C. (1981). A note on determining significant discrepancies among category scores on Bannatyne's regrouping of WISC-R subtests. *Journal of Learning Disabilities*, 14(8), 468-469.
- Reynolds, C., & Gutkin, T. (1980). Stability of the WISC-R factor structure across sex at two age levels. *Journal of Clinical Psychology*, 36(3), 775.
- Ryckman, D. (1981). Searching for a WISC-R profile for learning disabled children: An inappropriate task? *Journal of Learning Disabilities*, 14(9), 507-526.
- Salvia, J., & Ysseldyke, J. (1981). *Assessment in special education and remedial education* (2nd Ed.). Dallas, TX: Houghton-Mifflin.
- Sattler, J. (1982). *Assessment of children's intelligence and special abilities* (2nd Ed.). Boston, MA: Allyn & Bacon.
- Segel, D. (1934). *Differential diagnosis*. Baltimore, MD: Warwick & York.
- Smith, G. (1970). *A simplified guide to statistics for psychology and education* (4th ed.). New York, NY: Holt, Rinehart & Winston.
- Tew, B. (1977). Spina Bifida children's scores on the WISC. *Perceptual & Motor Skills*, 44(2), 381-382.
- Thompson, A. (1981). Some counter thinking about learning disabilities. *Journal of Learning Disabilities*, 14(7), 394-396.
- Thompson, R. (1980). Diagnostic utility of WISC-R measures with children referred to a developmental evaluation center. *Journal of Consulting & Clinical Psychology*, 48(4), 440-447.
- Tierney, R., & Ames, W. (1978). An examination of the diagnostic claims of the Illinois Test of Psycholinguistic Abilities. *Journal of Learning Disabilities*, 11(9), 586-590.
- Vance, H., & Singer, M. (1979). Recategorization of the WISC-R subtest scaled scores for learning disabled children. *Journal of Learning Disabilities*, 12(8), 487-490.
- Vance, H., Wallbrown, F., & Blaha, J. (1978). Determining WISC-R profiles for reading disabled children. *Journal of Learning Disabilities*, 11(1), 55-58.
- Van Hagen, J., & Kaufman, A. (1975). Factor analysis of the WISC-R for a group of mentally retarded children and adolescents. *Journal of Consulting & Clinical Psychology*, 43, 661-667.
- Wallbrown, F., Vance, H., & Blaha, J. (1979). Developing remedial hypotheses from ability (WISC-R) profiles. *Journal of Learning Disabilities*, 12(1), 59-63.
- Webster, R., & Lafayette, A. (1980). Distinguishing among three subgroups of handicapped students using Bannatyne's recategorization of the WISC-R. *Journal of Educational Research*, 73(4), 237-240.
- Wechsler, D. (1974). *WISC-R manual*. New York, NY: Psychological Corporation.

PETER EVANS

NORMA MICKELSON

and

NELSON C. SMITH

University of Victoria

# Computer-Assisted Instruction in College Composition

*This research investigated the use of microcomputers in the development of prose writing ability of university students. A computer-assisted language program employing the Apple II computer in a tutorial function was used in the study. The program presented information, responded to answers, and evaluated student input with a programmed response.*

*Two groups of 25 freshman students enrolled in basic composition courses at the University of Victoria were used as subjects. Student performance was measured through pre- and post-tests of sentence structure, error analyses (tabulated by the computer), and an examination of sentence types used in written compositions.*

*The results indicated few significant differences in performance between control and experimental groups. All differences, however, favored the experimental group. It was concluded that computer-assisted instruction appears to be at least as effective as regular instruction. This suggests that computer-assisted instruction can be used to supplement face-to-face teaching, thus freeing the instructor for individual attention to student needs.*

## Introduction

The objective of this research was to investigate the use of microcomputers in the development of prose writing ability at the university level. The research was

---

Peter Evans is a Professor of Language Arts at the University of Victoria. He has published numerous articles on aspects of language learning as well as collaborated in the writing of student texts for the junior high school reading program.

Norma Mickelson is a Professor of Education at the University of Victoria with extensive experience in public schools, both as a teacher and as a supervisor of instruction. The current focus of her work is in the application of learning principles to classroom practice.

Nelson C. Smith, Associate Professor of English at the University of Victoria, has been serving as Director of First Year English and the Language Programme for the past two years.

heuristic in nature, its purpose being to examine the potential of computer technology in the development of student writing achievement and to ascertain directions for future research and application.

The computer-assisted language program used in this study was developed to function as a "tutor." Apple II discs were used. The program presented information, responded to answers, and evaluated student input with a programmed response. While it was a right/wrong format and thus did not evaluate the quality of the errors, it did make the student aware of mistakes and also supplied correct responses. Error detection by the computer enabled the student to assume responsibility for analyzing and correcting mistakes. The pupils, however, if they so desired, could refer to the instructor for discussion and explanation of errors. This approach seemed particularly appropriate at the college level where student-directed learning is one goal of instruction and where remedial activities appear to be necessary for improvement of writing skills.

Since student attitude to writing is a variable which may affect composition, the researchers used the Writing Apprehension Measure (WAM) (Daly & Miller, 1975b) to assess student attitudes. Daly hypothesized that high writing apprehension scores correlate negatively with success in writing and also predicted that females would score lower in writing apprehension than would males. The Writing Apprehension Measure is a 26-item Likert scale with a range of five responses. Items represent anxiety about writing. The scale identifies highly apprehensive students who find writing unrewarding and anxiety-producing at one end of the continuum, and at the other, less apprehensive individuals who are positive and confident in their attitudes towards writing.

In summary, then, the study addressed three basic questions:

1. Does computer-assisted instruction affect writing performance?
2. What are student attitudes to computer-based instruction?
3. Does attitude towards writing affect performance?

### *Review of the Literature*

Readings in the application and implementation of the microcomputer in the language arts consist mainly of position papers and theoretical constructs. The scarcity of empirical studies reflects the short history of computers in the classroom and the more frequent application of computer-assisted instruction to the mathematics and science areas.

Arguments for and against computer-assisted language learning were articulated by Spolsky (1970) in a discussion of the nature of language learning and the theoretical basis of grammar teaching using a computer. Spolsky argued that the computer was a valuable tool because it forced the student-operator to develop a rationale for his problem-solving, thereby enhancing his achievement.

Gagné (in Stotsky, 1982) saw the computer as having potential for the teaching of basic skills in language arts classes. He discussed the need for bringing basic skills to the point of automaticity in order to supplement the limited capacity of short-term memory in problem-solving. He pointed out the utility of the computer in providing consistency in the drill and practice techniques needed to bring subordinate skills to a level of automaticity.

Many authorities in language education have stressed the need for instruction in grammar, spelling, word definition, and sentence construction to ensure a reason-



able level of composition ability and literacy. Others have pointed to the potential of the computer to motivate students and make routine material interesting. The importance of increased motivation was specifically mentioned by Holmes and Kidd (1982). The tutorial method of computer-assisted instruction seemed to them to be an effective use of the microcomputer's unique characteristics. Interaction between the learner and the material is possible when "branching" is used to allow for individual pacing and reinforcement (Holmes & Kidd, 1982).

Boyd, Keller, and Kenner (1982) felt that grammatical constructions lend themselves readily to effective computer-assisted instruction by means of multiple choice items.

The teaching of grammar and its effects on writing ability continue to be discussed in the literature. Stotsky (1975) pointed to the failure of past studies of grammar usage to test for skills taught. In a review of the research she included the work of Mellon (1982) and O'Hare (1973) which showed an improvement in T-unit measures (Loban, 1963) and an increase in variety of sentence structures as a result of instruction.

Although a consensus of opinion in the literature suggests that the teaching of grammar in a traditional sense does not necessarily result in improvement of writing skills, it seems possible that using a computer as a novel approach to teaching basic grammar skills might reverse this situation.

### *Design*

#### *The Research Model*

A review of the literature reveals two fundamental factors which impinge upon the present research. Firstly, there are mixed opinions and conflicting evidence regarding the impact upon composition of the teaching of fundamental grammatical concepts such as sentence types and coordination and subordination of sentence elements. Secondly, there is little evidence of the efficacy of computer-based instruction upon the learning of grammatical concepts. Added to these two factors are concerns about other antecedent conditions surrounding the teaching and learning of composition skills. These include student attitudes to writing and the differential influence of sex upon performance.

In order to examine the interrelationships of these factors, this study was organized to examine the instrumental effects of learning fundamental grammar concepts upon composition. Further, two approaches to the development of these concepts were compared. Finally, the effects of attitude toward writing and the sex influence on performance were investigated. The research design employed was the Static Group Comparison Model (Campbell & Stanley, 1963, p. 8). This model permits comparison of the performance of two intact groups, one of which receives the experimental treatment.

#### *Subjects and Procedures*

Two groups of 25 freshman students enrolled in basic composition courses at the University of Victoria were used as subjects. All freshman students had been randomly assigned to 16 sections of the course. The two groups used in the research were selected on the basis of scheduling availability. Through attrition (dropping of the course) four were lost from one group and three from the other.

The control group (RI) was given regular classroom instruction in sentence structure. The experimental group (CAI) was given similar instruction using an

Apple II computer for three one-hour sessions. Both instructional programs presented, described, and classified sentence types. Clauses were identified and bracketed; sentence errors and effective grammatical devices were identified by example and definition. For the control group (RI), these definitions were presented by the teacher in traditional classroom instruction. Traditional classroom instruction in this study involved the teaching of basic writing skills including the knowledge of grammatical structures and rhetorical devices through face-to-face presentation. Concepts were presented and defined. This activity was followed by student application in assigned writing activities. For the experimental group, an Apple II Disc (tutorial) led the student through a series of activities designed to cover the same material as in-class instruction. Both programs included simple sentences, complex sentences, compound sentences, effective use of clauses, subordination, coordination, parallelism, and sentence errors. The student's task was to select ways to correct faulty sentences.

### *Instrumentation*

Student performance was measured as indicated in Table 1.

Variables 1 and 2, the Objective Test, measured the students' knowledge of the specific concepts taught only in the computer tutorial program. This 40-item exercise was a pencil and paper test, and the pretest was used as the covariate in the statistical analysis.

Variables 3 and 4, the Essay Marks, were a general assessment of the quality of the students' writing. The relationship between the learning of grammatical concepts and the quality of writing was examined. Variables 5 to 14 and 23 to 28, sentence types and rhetorical devices, were used as a specific analysis of the degree to which the treatment and control learning activities influenced the use of these structures in writing. Variables 5-14 were pre-treatment results and variables 23-28 were post-treatment results.

Variables 17 and 18, Error Analysis, were the results of the computer tabulation of student performance on the computer tutorial activities. These variables indicated the degree of difficulty students had with the lesson material.

Variables 20, 21 and 22, Course Marks and Marks on Sentence Structure in the Final Exam, were performance values assigned by instructors for work done in the course. These evaluations were examined to determine the relationships between specific computer learning activities and general performance in the course.

Finally, student attitudes were measured by Daly's Writing Apprehension Measure and the results correlated with the grammar and composition scores.

### *Analysis*

An analysis of covariance was used to analyze the data with the pre-treatment grammar scores as the covariate. The level of significance selected for the study was .05.

## *Results*

### *A. Experimental Results*

Results indicate few significance differences in performance between control and experimental groups. Both groups, however, improved over time. Computer-assisted instruction, therefore, appears to be no less effective than regular instruction.

TABLE 1  
ANALYSIS OF COVARIANCE

| Variable Number | Variable Name                                 | Male/<br>Female | CAI/<br>RI | WAM Score<br>Low/High |
|-----------------|---|-----------------|------------|-----------------------|
| VO1             | Objective Pre-Test                            | NS              | NS         | NS                    |
| VO2             | Objective Post-Test                           | NS              | NS         | S                     |
| VO3             | Essay Mark: Pre-Experimental                  | NS              | NS         | NS                    |
| VO4             | Essay Mark: Post-Experimental                 | S               | NS         | NS                    |
| VO5             | Essay—No. of Simple Sentences: Pre-Exp.       | NS              | NS         | NS                    |
| VO6             | Essay—No. of Compound Sentences:<br>Pre-Exp.  | NS              | NS         | NS                    |
| VO7             | Essay—No. of Complex Sentences: Pre-Exp.      | NS              | S          | NS                    |
| VO8             | Essay—Total No. of Sentences: Pre-Exp.        | NS              | NS         | NS                    |
| V13             | Essay—No. of Sentence Errors: Pre-Exp.        | NS              | S          | S                     |
| V14             | Essay—No. of Effective Devices: Pre-Exp.      | NS              | NS         | NS                    |
| V17             | Error Analysis I CAI                          | NS              |            | NS                    |
| V18             | Error Analysis II CAI                         | NS              |            | S                     |
| V20             | Final Course Mark                             | NS              | NS         | NS                    |
| V21             | Mark on Exam Sentence Structure               | NS              | NS         | NS                    |
| V22             | Final Exam Mark                               | NS              | NS         | NS                    |
| V23             | Essay—No. of Simple Sentences: Post-Exp.      | NS              | S          | NS                    |
| V24             | Essay—No. of Compound Sentences:<br>Post-Exp. | NS              | NS         | NS                    |
| V25             | Essay—No. of Complex Sentences: Post-Exp.     | NS              | NS         | NS                    |
| V26             | Essay—Total No. of Sentences: Post-Exp.       | NS              | NS         | NS                    |
| V27             | Essay—No. of Sentence Errors: Post-Exp.       | NS              | NS         | NS                    |
| V28             | Essay—No. of Effective Devices: Post-Exp.     | NS              | NS         | NS                    |

NS = No Significance  
S = Significant Difference: See Table 2

An examination of the data reveals the following:

1. On the post-treatment objective test, the low-apprehensive group scored higher than the high-apprehensive group (Table 2: V02).
2. The mean percentage for females in the post-treatment essay was higher than for males.
3. The CAI group used more complex sentences than the RI group in the pre-treatment compositions. This difference between groups was not present in the essays written after treatment.
4. The CAI group made more sentence errors (comma splice, ineffective subordination, ineffective coordination and parallelism) than the RI group in the pre-treatment essays; post-test differences were not significant. The group that scored low on the Writing Apprehension Measure made more sentence errors in the pre-treatment essay test than did the high-scoring WAM group. On the post-treatment essay test the low-scoring group made fewer errors than did the higher-scoring WAM group.



- 5. CAI subjects having a low WAM score made more errors in the computer program than did those having a high WAM score.
- 6. The CAI group used fewer simple sentences in the post-treatment essay than did the RI group.

TABLE 2  
SIGNIFICANCE DIFFERENCES

| Variable  | Male | Female | CAI  | RI   | Signif.<br>of F | WAM Score |       | Signif.<br>of F |
|---|------|--------|------|------|-----------------|-----------|-------|-----------------|
|   |      |        |      |      |                 | Low       | High  |                 |
| VO2<br>Objective Post-Test                                    | —    | —      | —    | —    | —               | 25.78     | 21.73 | .033            |
| VO4<br>Essay Mark:<br>Post-Experimental                       | 64.5 | 69.35  | —    | —    | .025            | —         | —     | —               |
| VO7<br>Essay—No. of<br>Sentence Errors:<br>Pre-Experimental   | —    | —      | 9.10 | 5.25 | 0.11            | —         | —     | —               |
| V13<br>Essay—No. of<br>Sentence Errors:<br>Pre-Experimental   | —    | —      | 5.48 | 2.75 | .020            | 6.18      | 2.92  | .048            |
| V18<br>Error Analysis II CAI                                  | —    | —      | —    | —    | —               | 62.7      | 50.0  | .041            |
| V23<br>Essay—No. of<br>Simple Sentences:<br>Post-Experimental | —    | —      | 5.64 | 8.65 | .037            | —         | —     | —               |

TABLE 3  
OBJECTIVE TEST MEAN SCORES

|     | Pre-Test | Post-Test |
|-----|----------|-----------|
| CAI | 16.9     | 21.8      |
| RI  | 16.4     | 22.2      |

## B. Anecdotal Results

In analyzing student reactions (as recorded on individual log sheets) to the computer-assisted instruction, the researchers noted that students liked the individual attention provided, the chance to set their own pace, and the immediate feedback given in the program. Mentioned also was the lack of distraction. Many comments suggested the need for frequent loops to allow for referral back in the program for information, the usefulness of frequent repetition to provide reinforcement, and the desirability of being able to start or stop midway through the program. One student with previous experience with computers recommended more incentives or rewards to eliminate boredom. Students also expressed interest in knowing their scores because they thought this would help them evaluate their progress. Two subjects mentioned frustration when they did not understand why their answers were incorrect. In general, females appeared to be more apprehensive about using computers than did males.

## Discussion

While at first glance there appear to be few differences in effect between regular and computer-assisted instruction, a detailed analysis of the data reveals noteworthy differences favoring the treatment (CAI) group. These differences are discussed below under each of the three basic questions considered in the study.

### A. Does computer-assisted instruction affect writing performance?

1. Results indicate a significant difference between the CAI group and the RI group in use of complex sentences in the pre-treatment essay. This difference was not noted in the post-test essay. While this may appear to be a reduction in the quality of performance, an examination of the essays revealed that this result indicated, in fact, a reduction of run-on sentences in the compositions of the CAI group. It should be noted that in the original design of the study, the analysis did not include run-on sentences. These were categorized as complex sentences. The results appear to indicate increased awareness by the students of the importance of avoiding run-on sentences.
2. The result showing a significance difference in essay writing following instruction in favor of females in both groups suggests that in both computer-assisted instruction and regular instruction, females improved significantly. This result would tend to contradict the pervasive assumption that women do not benefit from technological approaches to learning.
3. Significant differences in the number of simple sentences used in the post-test favoring the regular instruction group suggest that students showed less improvement in the use of a variety of sentence types than did those in the computer-assisted group. Thus the tutorial program may have left a more powerful impression on the learners as to the variety of sentence forms and structures available for use in written composition than did regular classroom instruction.

### B. What are students' attitudes to computer-based program instruction?

1. In general, most students reacted positively to computer-assisted instruction as evidenced by comments such as, "It's fun" and "I can set my own pace."

2. Students noted the need for an explanation of incorrect answers in the computer tutorial. This observation reinforces the importance of teacher participation in learning and suggests that computers might best be used as an adjunct to regular classroom instruction.
3. Many of the disadvantages of computer-assisted instruction expressed by students could be eliminated in more sophisticated programs with frequent loops for optional reviewing, reinforcement, or further explanation. A summary of the student's correct and incorrect answers could also be provided with areas of difficulty pinpointed.

### *C. Does attitude towards writing affect performance?*

In the post-treatment objective grammar test, students with low WAM scores performed better than did those who were more apprehensive, a finding consistent with Daly's hypothesis. In essay writing, high apprehensive students appeared to be more concerned about mechanics than did low apprehensive students. Low apprehensive students appeared to profit more from instruction as indicated by sentence error data. The result showing that low apprehensive students made more errors in the computer tutorial suggests that they appear to be willing to risk making errors to a greater degree than do students who are highly apprehensive. More research is needed in this area.

### *Conclusions and Recommendations*

The results of this study suggest that computer-assisted instruction appears to be no less effective than regular classroom teaching of sentence mechanics. The fact that there were no measurable differences in performance of students on most variables indicates that computer-assisted instruction may produce results as good as face-to-face instruction. In no case did subjects who were exposed to the microcomputer in place of regular instruction achieve less well than did those who attended regular classes. In fact, there appeared to be several important benefits for those who used the computer. For example, students developed both greater ability to use variety in sentence structure and accuracy in construction of sentences. Outcomes such as these are considered to be fundamental to a program of English composition instruction.

Computer-assisted programs of instruction in fundamental composition skills appear to be both effective and viable. Such programs take less teacher time than regular classroom instruction and thus might effectively supplement and vary approaches taken in the teaching of English. The more factual and technical aspects of course content with the focus on student responsibilities for learning, for example, could be presented by computers, while the application of such content might be handled by human instructors. Remedial teaching could well be handled, at least in part, using computer-assisted instruction.

Specific recommendations arising out of the study are as follows:

1. There were certain problems in logistics during the experimental program. These involved the availability of microcomputers and the scheduling of appropriate practice sessions. The study should be replicated in such a manner as to obviate these problems.
2. The identification of errors and sentence structures in essays should be more rigidly controlled.



3. Students in first-year English should be given an opportunity to supplement the learning of fundamental composition skills through computer-assisted instruction.
4. Females, in particular, should be encouraged to participate in computer-aided courses of instruction.
5. The writing apprehension level of students might be investigated further as a criterion in planning learning activities.
6. Studies might be conducted to determine how increased exposure to computer programs affects learning rate.

### References

- Adams, E. N., Morrison, H. W., & Reddy, J. M. (1968). Conversations with a computer as a technique of language instruction. *The Modern Language Journal*, 42, 3-15.
- Boyd, G., Keller, A., & Kenner, R. (1982). Remedial and second language English teaching using computer-assisted learning. *Computers and Education*, 6, 105-112.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.
- Daly, J. A. (1977). The effects of writing apprehension on message encoding. *Journalism Quarterly*, 54, 566-572.
- Daly, J. A., & Miller, M. D. (1975a). Apprehensions of writing as a predictor of message intensity. *Journal of Psychology*, 89, 556-572.
- Daly, J. A., & Miller, M. D. (1975b). The development of a measure of writing apprehension. *Research in the Teaching of English*, 9, 242-249.
- Daly, J. A., & Shamo, J. (1975). Writing apprehension and occupational choice. *Journal of Occupational Psychology*, 49, 55-56.
- Holmes, G., & Kidd, M. E. (1982). Second-language learning and computers. *The Canadian Modern Language Review*, 3, 503-516.
- Loban, W. (1963). *The language of elementary school children*. Urbana, IL: (NCTE Research Report No. 1).
- Mellon, J. C. (1982). *Transformational sentence-combining: A method for enhancing the development of syntactic fluency in English composition*. Cambridge: Harvard University.
- O'Hare, F. (1973). *Sentence-combining: Improving student writing without formal grammar instruction*. Urbana, IL: (NCTE Research Report No. 15).
- Spolsky, B. (1970). The language teacher in the computer age. In R. Lugton (Ed.), *Preparing the E.F.L. teacher*. Philadelphia: The Center for Curriculum Development.
- Stotsky, S. (1975). Sentence-combining as a curricular activity: Its effect on written language development and reading comprehension. *Research in the Teaching of English*, 9, 30-71.
- Stotsky, S. (1982). Developments in learning psychology: Implications for instructional design: Effects of computer technology on instructional design and development. An interview with Robert M. Gagné, in *Educational Technology*.

WAYNE K. HOY

*Rutgers University*

and

DAVID A. SOUSA

*West Orange Public Schools*

## Delegation: The Neglected Aspect of Participation in Decision Making<sup>1</sup>

*Data were collected from the principals and teachers in 55 secondary schools. Using the Aston methodology, an index of shared decision making was developed, and three hypotheses were derived and tested. All were supported. Principals who shared decision making with subordinates had more loyal teachers, had more satisfied teachers, and had teachers who described the authority structure as less rigid than principals who did not share decision making. The research underscored an important theoretical distinction between sharing decision making authority in terms of delegating and sharing in terms of joint decision making.*

The issue of participation in decision making remains a popular topic in organizational analysis; in fact, subordinate participation is a basic tenet of human relations theory (Likert, 1967; Strauss & Rosenstein, 1970). Despite its common usage, there is relatively little consensus on the nature and meaning of participation. Schregle (1970, p. 117) cogently comments, "Workers' participation has become a magic word in many countries. Yet almost everyone who employs the term thinks of something different." Although Shregle's observation seems a bit strong, participation in decision making clearly suffers from a variety of uses (Dachler & Wilpert, 1978). For example, Likert (1967) views participation as a management style; Leavitt (1965) and Tannenbaum (1974) see it as equalization of power or sharing power; and Davis (1967) relates it to group involvement and group decision making. Moreover, some theorists distinguish between participation in decision making and

---

Wayne K. Hoy is Professor and Associate Dean for Academic Affairs, Graduate School of Education, Rutgers University. He is coauthor of *Educational Administration: Theory, Research and Practice*.

David A. Sousa is Director of Science, West Orange Public Schools, West Orange, New Jersey.

delegation of decision making (Davis, 1963; Locke & Schweiger 1979); others do not (Lowin, 1968; Sashkin, 1976).

In the present investigation the concern was with examining one aspect of participation in decision making—the extent to which principals delegate decisions to subordinates. The approach was first to develop an index of the principal's propensity to share important decisions with subordinates and then to examine the relationships of that propensity with teachers' loyalty to principal, teachers' satisfaction with their jobs and teachers' perception of the degree of hierarchical control.

### *Conceptual Perspective*

Four specific concepts provided the perspective employed in the study; each is identified and described below.

*Participation in Decision Making* is a broad term often used to refer either to joint decision making or to the process of delegating specific decisions to subordinates. The process of delegating decision authority to subordinates, however, is not the same as jointly arriving at a decision. In joint decision making the participants are involved in the *process* of making decisions; they share a common effort. *Delegation of decision making* is the entrusting of authority to others; an administrator assigns specific decisions to others, usually subordinates. The subordinates are then free to make their decisions without further consultation with the supervisor. Many studies have centered on participation in decision making but few in recent years have examined the benefits of delegated decision making.

Principals who delegate recognize that they are yielding some of their authority and that, for the most part, they are bound by the subordinates' decisions. What are the benefits? Does such a practice improve job satisfaction among subordinates? Does it increase subordinate loyalty? Does it alter teachers' perceptions of the authority structure of the school? These were the basic questions that guided the empirical phase of this study.

*Hierarchy of Authority* is the extent to which the focus of decision making is prestructured by the formal authority system. Organizations vary in the degree to which decisions must be first approved by superiors before any action can be taken. School organizations in which teachers are given little or no freedom to accomplish assigned tasks unimpeded by their superiors have a rigid hierarchy of authority. What we are calling hierarchy of authority has also been termed tightness or closeness of supervision by other researchers (Blau & Scott, 1962, pp. 145-150).

*Job Satisfaction* reflects a sense of gratification with one's position of employment in the organization; it is the extent to which an individual is satisfied with such things as authority, relative position compared to other workers, opportunities for professional growth, recognition and acceptance granted by superiors, and the degree to which the job is consistent with one's career expectations. Aiken and Hage (1966) call the lack of satisfaction with one's job alienation from work; this study uses the positive term, job satisfaction, because it more accurately reflects the meaning of the operational definition and it avoids the questionable comparison of the concept with Marx's alienation from the process of production (Schacht, 1970).

*Subordinate Loyalty to Superior* was first introduced by Blau and Scott (1962) as an important variable mediating work relations in hierarchical organizations, but despite its apparent theoretical significance, they simply noted that superiors who command loyalty are "liked," "accepted," and "respected." Murray and Corenblum (1966) defined the term more rigorously, arguing that, in addition to the affect-



based definition of Blau and Scott (1962), at least two other dimensions of subordinate loyalty—cognitive and behavioral—should be considered. Subordinates who are cognitively loyal manifest an unquestioning faith and trust in the superior, and those who are behaviorally loyal show an actual or expressed willingness to follow when the superior accepts a new position. Subsequent research in schools (Hoy & Rees, 1974; Hoy & Williams, 1971) demonstrated that the three dimensions of subordinate loyalty vary concomitantly; hence, *loyalty to the principal* is the extent to which the principal is liked, respected, accepted, trusted and followed by teachers.

The research, then, examined the relationships between delegated decision making and the three concepts of (1) loyalty, (2) job satisfaction, and (3) hierarchy of authority.

### *Some Hypotheses*

Principals are required to make many decisions in the course of their work. Some decisions are trivial while others have a major impact on the operation of the school. Principals have devised various methods for dealing with decision making. Those who make decisions without input from other parts of the organization have been referred to as authoritarian. At the other extreme are those who seek extensive input before deciding. These are often referred to as democratic or participative administrators. In his study, Cross (1980) noted that principals who sought input relied on one or more of the following four sources of information when making decisions: subordinates, extraordinates (people not directly affiliated with the school), superordinates, and peers. He found that although principals rarely consulted superordinates, they were strongly influenced in their decisions by subordinates.

Teachers in recent years have requested greater participation in decisions that affect them and in which they have some expertise. The delegation of such decisions by principals to subordinates is likely to influence their perceptions of the control structure of the school. In a study of social welfare agencies, Aiken and Hage (1966) found a substantial negative correlation between the degree of subordinate participation in decision making and extent of hierarchical control. Moreover, Hoy and his colleagues (1983) report the same results in their comprehensive study of principals and teachers in 41 high schools in New Jersey; the greater the degree of teacher participation in decision making, the less the sense of hierarchy of authority. In both studies, however, participation in decision making was defined in broad terms. In each case subordinates were asked such questions as, "How frequently do you participate in the decision to hire new staff?" and "How frequently do you participate in decisions on the adoption of new policy?" It is unclear, however, whether participation meant the act of being consulted by superiors, or jointly making the decisions with superiors, or making decisions on their own that had been delegated to them by superiors. Delegating the decision to subordinates has the potential to maximize participation; the principal entrusts the decisions to subordinates. The extent to which such delegation occurs seems likely to affect the teachers' perception of the formal authority structure of the school; consequently it was hypothesized that:

- H.1. The greater the propensity of the principal to delegate decision making, the less hierarchy of authority perceived by teachers.

Teachers do not want to be involved in all kinds of decisions. Kunz and Hoy (1976) found that school decisions were characterized by a "zone of acceptance,"

and that teachers willingly left decisions within this zone to administrators. In an extensive three-year study, Crockenberg and Clark (1979) were able to get active teacher participation only after formal mechanisms were established through faculty constitutions and faculty councils. Moreover, they found that the decision-making areas of particular concern to teachers were distinctly instructional rather than administrative. Duke, Showers, and Imber (1980) found that most teachers were willing to participate in decision making only when such involvement meant actual influence over decisions, and that these decisions had great potential for improving classroom life and student outcomes. Healey (1980), in a survey of 150 elementary principals in Illinois, also found that the more the decisional task related to classroom instruction, the more faculty involvement in decision making occurred.

Job satisfaction reflects a sense of satisfaction with one's career, professional development, authority, position, and ability to fulfill professional responsibilities. Aiken and Hage (1966) referred to such lack of satisfaction with one's job as work alienation, and several studies (Aiken & Hage, 1966; Hoy et al., 1983) found that the lack of participation by subordinates in organizational decision making was strongly related to their alienation from work. Conway's (1976) research showed that secondary school teachers were more satisfied with their jobs if they participated in and influenced decisions affecting instruction, and if such participation was not excessive. The Isherwood and Taylor (1978) study of 48 secondary schools in Quebec found that teacher participation in decision making through school councils had a positive effect on their job satisfaction. In their extensive review of studies on decision making in non-educational organizations, Locke and Schweiger (1979) concluded that a majority of their studies supported the premise that increased participation in decision making resulted in increased job satisfaction. In schools, if teachers are granted authority to make important decisions, then they should feel a greater sense of self-attainment and influence, thereby increasing their satisfaction with their jobs; hence, it was hypothesized that:

H.2. The greater the propensity of the principal to delegate decision making, the greater the job satisfaction of teachers.

Many principals have only the power and authority of their office. In a real sense they are merely officers or bureaucrats, not leaders. How can principals increase the base of their authority? The informal organization of teachers is an important source of authority, and principals need to find ways to tap the informal authority network. Where formal authority is legitimated by the values that have become routinized in legal contracts, informal authority is legitimated by the common values and sentiments that emerge in the work group, *particularly by the loyalty the superior commands from group members* (Blau & Scott, 1962, p. 144). The importance of teacher loyalty to principals seems clear. Those principals who can command loyalty from teachers have an advantage in enlarging their authority base by establishing effective informal authority over teachers (Hoy & Rees, 1974, p. 269).

Several studies have shown that rigid control was negatively related to subordinate loyalty to the supervisor. Blau and Scott (1962) provided original evidence that tight control in social welfare agencies was negatively related to commanding subordinate loyalty. Similarly, two studies in secondary schools demonstrated that the more authoritarian the behavior of the principal, the less the teacher loyalty to the principal (Hoy & Rees, 1974; Hoy, Tarter, & Forsyth, 1978). Moreover, in a study of secondary school principals, a global measure of participa-



tion in decision making was found to be positively related to teacher loyalty to the principal (Hoy, Newland, & Blazovsky, 1977). When a principal voluntarily delegates important decisions to a teacher or group of teachers, it reflects the principal's confidence in the ability of teachers to reach acceptable, reasonable, and useful decisions. Such displays of confidence and trust should increase the sense of loyalty teachers feel toward the principal; therefore, it was hypothesized that:

H.3. The greater the propensity of the principal to share decision making, the greater the degree of teacher loyalty to the principal.

### *Design and Method*

In order to test the hypotheses of this study, an index of propensity to delegate decision making was developed for a sample of secondary school principals; and the perceptions of the teachers in each school were analyzed.

#### *Sample*

Data were collected on 55 public senior high schools located in 11 counties in New Jersey. In selecting the schools, an attempt was made to obtain a variety of community types including urban, suburban, and rural areas. The size of the high schools in the sample ranged from 339 to 2400 students (with a mean of 1251 students), and the full-time professional staff ranged from 34.5 to 168 (with a mean of 97.0 professionals per school). Although no claim is made that the sample of schools was random, it seems to be a representative sample of New Jersey secondary schools, especially with respect to size. If any type of school was underrepresented, it was probably the large urban-core high school—only two such schools were included in the sample.

#### *Independent Variable—Propensity to Delegate*

In conducting this study, it was necessary to select a set of decisions that were typical for public school settings, that were relevant and important, and that discriminated among schools in terms of the principal's decision-making authority. The Aston approach developed by Pugh and Hickson (1978) and their colleagues was deemed appropriate for the task.<sup>2</sup> This approach analyzes the structure of organizations through interviews with the chief executives. The interview schedule, designed to gather more objective data about the organization, was administered to each principal. One portion of the schedule contained a series of school-related decisions that described the principal's autonomy. In effect, the autonomy items measured the number of decisions that were made by building principals without the immediate approval of their superordinates. Those decisions that were autonomous to principals represented the maximum number of decisions that principals in each building could legitimately delegate to subordinates. Specifically, principals were asked to identify the level within the school system at which 22 different decisions were made. The six decision making levels from which they could select were: teacher, department head, assistant principal, principal, superintendent, and local board of education. It was explained that decision making authority meant action could be taken on the decision, even though the decision might be subject to routine ratification at a later time. This meant that, although others might ratify the decision, its intentions would not be altered.



The principals' responses were subjected to item analysis, and the ten decision questions listed below were found to discriminate among the schools.

Who decides:

1. The number of department heads in the schools?
2. The appointment of a teacher?
3. The amount of money to be allocated to each department?
4. How unallocated funds will be spent?
5. Whether a new course or subject will be introduced?
6. Whether a new program will be introduced?
7. The procedures for purchasing materials for the school?
8. Which suppliers or materials will be used?
9. To create a new department?
10. To create a new teaching or administrative position?

By identifying the level at which decisions were made for each school, the principals provided a means for investigating the extent to which they delegated their decision-making authority. Specifically, the comparison was made by establishing a ratio in which the numerator represented the number of decisions made by the subordinates (teachers, department heads, or assistant principals) in the school, and the denominator represented the total number of within-school decisions made. This ratio was called the delegation of decision making index and was expressed by the equation:

$$\frac{\text{Number of subordinate decisions}}{\text{Total number of within-school decisions}} = \text{Delegation of decision making index}$$

The index had a range of 0 to 1.00. The more decisions that were made by the subordinates, the higher the value of the index and the greater the principal's propensity to delegate decision making.

### *Dependent Variables*

During a regular faculty meeting, one-quarter of the faculty, selected at random, responded to measures of the dependent variables—hierarchy of authority, job satisfaction, and loyalty to the principal—while the rest of the faculty completed other research measures not used in this study. Virtually all teachers selected completed the research instruments. A total of 922 (96 percent) of the 965 forms completed was usable. Individual scores were combined into a mean score for each school; hence, the school scores were used in all hypothesis testing.

*Hierarchy of Authority* was measured with a fourteen-item index developed from the work of Hall (1963), McKay (1964), and Aiken and Hage (1966). Teachers were asked to respond to each item along a five-point response scale of "always," "often," "occasionally," "seldom," or "never" (always=5). Examples of items were: "Any decision I make has to have my superior's approval," "I have to ask the principal before I do almost anything," and "There can be little action until a supervisor approves a decision." The reliability for the measure was supported by an alpha coefficient of .89 for the sample. Construct validity of the measure has been demonstrated in a number of studies (Sousa & Hoy, 1981; Punch, 1969).

*Job Satisfaction* was operationalized using an eight-item index developed by Aiken and Hage (1966). Teachers were asked to answer the questions in terms of their degree of satisfaction along a four-point scale from 4 (very satisfied) to 1 (very

dissatisfied). Samples of questions included the following: "How satisfied are you that you have been given enough authority by administrators to do your job well?" "How satisfied are you with your present job when you compare it to similar positions?" and "How satisfied are you with your present job in light of career expectations?" Alpha coefficients for the index have been consistently about .80 (Hoy, Blazovsky, Newland, 1983); in the present study the alpha was .83. Likewise construct validity of the measure has been supported in a number of studies (Aiken & Hage, 1966; Hoy, Blazovsky, & Newland, 1983).

*Loyalty to the Principal* was measured with an eight-item index adapted by Hoy and Williams (1971) from the work of Murray and Corenblum (1966). The following questions are examples of the items to which teachers responded: "If you had a chance to teach for the same pay in the school under the direction of another principal, how would you feel about moving?" "Generally speaking, how much confidence and trust do you have in your principal?" "All in all, how satisfied are you with your principal?" Construct validity for the measure has been supported in several studies (Hoy, Tarter, & Forsyth, 1978; Hoy & Rees, 1974; Hoy & Williams, 1971). Reliability has been consistently high with alpha coefficients in the .90 range; in this study the alpha was .92.

Results

Although 19 principals in the present sample delegated some decisions to subordinates, 36 principals delegated no decision making authority to teachers in the ten important decision making areas. Thus, the independent variable was dichotomized into delegation or no delegation of decision making, and the three hypotheses guiding the study were tested using analysis of variance procedures.

TABLE 1  
SUMMARY AND ANALYSIS OF VARIANCE DATA: THE RELATIONSHIP OF  
HIERARCHY OF AUTHORITY AND PRINCIPALS' DELEGATED DECISION-MAKING

|                    | Delegation     |    | No Delegation |      |      |
|--------------------|----------------|----|---------------|------|------|
| Number             | 19             |    | 36            |      |      |
| Mean               | 2.84           |    | 3.20          |      |      |
| Standard Deviation | .31            |    | .25           |      |      |
| Source             | Sum of Squares | df | Mean Square   | F    | P    |
| Between Groups     | 1.621          | 1  | 1.621         | 18.9 | <.01 |
| Within Groups      | 4.553          | 53 | 0.086         |      |      |
| Total              | 6.174          | 54 | 0.114         |      |      |

All three hypotheses were supported. Principals who delegated their decision-making authority had teachers who described the authority structure of the school as significantly less rigid than teachers who had principals who did not delegate ( $F=18.9, p<.01$ ). Likewise, principals with a propensity to delegate had teachers

who were significantly more satisfied with their job ( $F=10.4, p<.01$ ) and who were significantly more loyal ( $F=5.6, p<.05$ ) than principals who did not delegate. The analysis of variance data are summarized in Tables 1-3.

TABLE 2

SUMMARY AND ANALYSIS OF VARIANCE DATA: THE RELATIONSHIP OF  
JOB SATISFACTION AND PRINCIPALS' DELEGATED DECISION MAKING

|                    |                |            |             |               |      |
|--------------------|----------------|------------|-------------|---------------|------|
|                    |                | Delegation |             | No Delegation |      |
| Number             |                | 19         |             | 36            |      |
| Mean               |                | 2.93       |             | 2.71          |      |
| Standard Deviation |                | .23        |             | .25           |      |
| Source             | Sum of Squares | df         | Mean Square | F             | P    |
| Between Groups     | 0.606          | 1          | 1.606       | 10.4          | <.01 |
| Within Groups      | 3.088          | 53         | 0.058       |               |      |
| Total              | 3.694          | 54         | 0.069       |               |      |

TABLE 3

SUMMARY AND ANALYSIS OF VARIANCE DATA: RELATIONSHIP OF TEACHER  
LOYALTY AND PRINCIPALS' DELEGATED DECISION MAKING

|                    |                |            |             |               |      |
|--------------------|----------------|------------|-------------|---------------|------|
|                    |                | Delegation |             | No Delegation |      |
| Number             |                | 19         |             | 36            |      |
| Mean               |                | 3.57       |             | 3.28          |      |
| Standard Deviation |                | .48        |             | .43           |      |
| Source             | Sum of Squares | df         | Mean Square | F             | P    |
| Between Groups     | 1.100          | 1          | 1.100       | 5.6           | <.05 |
| Within Groups      | 10.460         | 53         | 0.197       |               |      |
| Total              | 11.560         | 54         | 0.214       |               |      |

*Summary and Discussion*

The findings in this study suggest that the willingness of principals to delegate important decisions to teachers (and other subordinates) can result in potential benefits for the school organization. First, teachers are less likely to perceive a



strong authority structure and thus may be more apt to identify with the goals and objectives of the school. Second, teachers gain a greater sense of job satisfaction, which tends to produce improved attitudes toward their work and the people with whom they work. Third, teachers are more inclined to exhibit loyalty to their principals.

The confirmation of all three hypotheses supports the validity of the index of delegation of decision making. The measure of propensity to delegate was determined by first establishing the important and specific internal decisions that the principal had the authority to make and then ascertaining which, if any, of those decisions were given to subordinates. The assumption was made that principals could delegate legitimately only decisions for which they had authority. Moreover, if principals shared important decisions, it was expected that teachers would view the hierarchy of authority as more flexible, an assumption that was supported by the findings. Further, we postulated that if teachers were granted authority to make important decisions, then they would feel a greater sense of self-attainment and influence thereby increasing their job satisfaction. Again, the findings supported the underlying theory.

A basic challenge facing all principals is to find ways to extend the scope of their influence over teachers beyond the narrow limits of formal authority flowing from their position. Barnard (1938, p. 170) has cogently suggested that only when the authority of leadership is combined with the authority of position will superiors be able to induce subordinates to comply with orders outside their bureaucratic zone of indifference. The informal organization is an important source of leadership, and principals who can command loyalty from teachers seem to have a distinct advantage in enlarging their authority over teachers (Hoy & Rees, 1974). The findings of this study suggest that principals who have the propensity to delegate decisions have a significant edge in commanding teacher loyalty. When principals entrust to teachers the authority to make important decisions, it seems likely teachers reciprocate by developing norms of allegiance, trust, and loyalty to the principal. Moreover, commanding subordinate loyalty has been linked with productivity. Blau and Scott (1962, p. 144) report that, "workers' productivity was largely associated with factors that were related to loyalty perhaps because supervisors who had won the loyal support of their subordinates were most successful in commanding willing compliance with their directives and in stimulating effort in their work group." Yet the results of this study show the principals are extremely reluctant to share their authority, with 36 of the 55 principals not delegating in the ten important areas.

The evidence, however, continues to grow that secondary school principals have significant authority to run their schools. For example, consistent with the findings of this study, two national studies of secondary school principals (Abramowitz & Tenenbaum, 1978; Brynes et al., 1978) show that the majority of principals select their own teachers, with only confirmation from the central office. Furthermore, most principals have major authority in nearly all matters that deal with the school's internal operation. Despite this extensive power of many principals, and despite the apparent benefits derived from entrusting some authority to subordinates, most principals in this sample (65 percent) had *no* propensity to delegate decisions to teachers in areas which teachers seemed to possess both expertise and a personal stake in the outcome. Perhaps principals consulted with teachers or jointly involved them in the process of making decisions, but most did not assign even such tasks as deciding to introduce a new course or program, or deciding

which suppliers or materials to use—let alone delegate decisions involving the selection of teachers or the internal allocation of funds.

Perhaps many principals typically are more concerned with their own position, status, and career than they are with the professional growth and development of subordinates. Perhaps some principals are too insecure to share power. Thirty-five percent of the principals in this study, however, did transfer significant decision making authority to their subordinates. Apparently, some principals do see the benefits. It could be hypothesized that principals who do delegate major decision making authority are more secure and more likely to view their subordinates as professionals than those who do not. Of course, the test of such hypotheses awaits further empirical investigation.

Clearly, some principals who did not delegate allowed subordinates to participate in joint decision making, but they reserved the final decisions for themselves. This fact points to a practical distinction between delegated and consultative decision making. Although many principals are willing to inform and to solicit information from subordinates on pending decisions, they do not delegate the decision. The result is that teachers begin to doubt the benefits of such participation and voice skepticism over whether their so-called involvement makes any difference (Duke et al., 1980). Apparently, teacher participation in decision making that lacks influence may be less beneficial than no teacher involvement at all.

This study also points to important theoretical distinctions among delegated, consultative, and joint decision making. Delegation of decision making maximizes participation; consultation minimizes participation. The term participation is simply too ambiguous and broad. The time has come to more clearly specify the meaning of involvement in decision making. The concept of delegation is a step in that direction. Moreover, the index developed in this research should provide researchers with a useful tool for the continued study of a concept that has both theoretical and practical significance.

#### Notes

1. An earlier version of this paper was presented at AERA in Montreal (1983).
2. The Aston approach was first modified for use in public schools by a group of researchers from the University of Alberta (Newberry, 1971; Kelsey, 1973; Holdaway, Newberry, Hickson, & Heron, 1975; Sackney, 1976).

#### References

- Abramowitz, S., & Tenenbaum, E. (1978). *High school '77*. Washington: National Institute of Education.
- Aiken, M., & Hage, J. (1966). Organizational alienation: A comparative analysis. *American Sociological Review*, 31, 497-507.
- Barnard, C. (1938). *The functions of the executive*. Cambridge, MA: Harvard University Press.
- Blau, P. M., & Scott, W. R. (1962). *Formal organizations*. New York: Chandler.
- Brynes, D. R. et al. (1978). *The senior high school principalship: The national survey (Vol. 1)*. Reston, VA: National Association of Secondary School Principals.
- Conway, J. A. (1976). Test of linearity between teachers' participation in decision making and their perceptions of their schools or organizations. *Administrative Science Quarterly*, 21, 130-139.



- Crockenberg, V., & Clark, W. W. (1979). Teachers' participation in school decision making: The San Jose teacher involvement project. *Phi Delta Kappan*, 61, 115-118.
- Cross, R. (1980). A description of decision-making patterns of school principals. *Journal of Educational Research*, 73, 154-159.
- Dachler, H. P., & Wilpert, B. (1978). Conceptual dimensions and boundaries of participation in organizations: A critical evaluation. *Administrative Science Quarterly*, 23, 1-39.
- Davis, K. (1963). The case for participative management. *Business Horizons*, 7, 45-50.
- Davis, K. (1967). *Human relations at work: The dynamics of organizational behavior*. New York: McGraw-Hill.
- Duke, D. L., Showers, B. K., & Imber, M. (1980). Teachers and shared decision making: The costs and benefits of involvement. *Educational Administration Quarterly*, 16, 93-106.
- Hall, R. H. (1963). The concept of bureaucracy: An empirical assessment. *American Journal of Sociology*, 63, 32-40.
- Healey, R. W. (1980). Principals' perceptions of faculty involvement in decision making in public elementary schools of Illinois. *Illinois School Research and Development*, 16, 91-97.
- Holdaway, E. A., Newberry, J. F., Hickson, D. J., & Heron, R. P. (1975). Dimensions of organizations in complex societies: The educational sector. *Administrative Science Quarterly*, 20, 37-59.
- Hoy, W. K., Blazovsky, R., & Newland, W. (1983). Bureaucracy and alienation: A comparative analysis. *Journal of Educational Administration*, 21, 109-210.
- Hoy, W. K., Newland, W., & Blazovsky, R. (1977). Subordinate loyalty to superior, esprit, and aspects of bureaucratic structure. *Educational Administration Quarterly*, 13, 71-85.
- Hoy, W. K., & Rees, R. (1974). Subordinate loyalty to immediate superior: A neglected concept in the study of educational administration. *Sociology of Education*, 47, 268-286.
- Hoy, W. K., Tarter, C. J., & Forsyth, P. (1978). Administrative behavior and subordinate loyalty: An empirical assessment. *Journal of Educational Administration*, 16, 29-38.
- Hoy, W. K., & Williams, L. B. (1971). Loyalty to immediate superior at alternate levels in public schools. *Educational Administration Quarterly*, 7, 1-21.
- Isherwood, G. B., & Taylor, R. M. (1978). Participatory decision making via school councils. *The High School Journal*, 255-270.
- Kelsey, J. G. T. (1973). *Conceptualization and instrumentation for the comparative study of secondary school structure and operation*. Doctoral dissertation, University of Alberta, Edmonton.
- Kunz, D. W., & Hoy, W. K. (1976). Leadership style of principals and the professional zone of acceptance of teachers. *Educational Administration Quarterly*, 12, 49-64.
- Leavitt, H. J. (1965). Applied organizational change in industry: Structural, technological and humanistic approaches. In J. G. March (Ed.), *Handbook of Organizations*. Chicago: Rand McNally.
- Likert, R. (1967). *The human organization: Its management and value*. New York: McGraw Hill.
- Locke, E. A., & Schweiger, D. M. (1979). Participation in decision making: One more look. *Research in Organizational Behavior*, 1, 265-339.
- Lowin, A. (1968). Participative decision making: A model, literature critique, and prescription for research. *Organizational Behavior and Human Performance*, 3, 68-106.
- McKay, D. A. (1964). An empirical study of bureaucratic dimensions and their relationship to other characteristics of school organizations. Unpublished doctoral dissertation, University of Alberta, Edmonton, Alberta.
- Murray, V. V., & Corenblum, A. F. (1966). Loyalty to immediate superior at alternate hierarchical levels in a bureaucracy. *American Journal of Sociology*, 22, 77-85.



- Newberry, J. F. (1971). *A comparative analysis of the organizational structures of selected post-secondary educational institutions*. Unpublished doctoral dissertation, University of Alberta, Edmonton.
- Pugh, D. S., & Hickson, D. J. (1978). *Organizational structure in its context*. Westmead, Farnborough, Hants, England: Saxon.
- Punch, K. F. (1969). Bureaucratic structure in schools: Towards redefinition and measurement. *Educational Administration Quarterly*, 5, 43-57.
- Sackney, L. E. (1971). *The relationship between organizational structure and behavior in secondary schools*. Unpublished doctoral dissertation, University of Alberta, Edmonton.
- Sashkin, M. (1976). Changing toward participative management approaches: A model and methods. *Academy of Management Review*, 1, 75-86.
- Schacht, R. (1970). *Alienation*. Garden City, New York: Doubleday.
- Schregle, J. (1970). Forms of participation in management. *Industrial Relations*, 9, 112-117.
- Sousa, D. A., & Hoy, W. K. (1981). Bureaucratic structure in schools: A refinement and synthesis in measurement. *Educational Administration Quarterly*, 17, 21-39.
- Strauss, G., & Rosenstein, E. (1970). Workers' participation: A critical view. *Industrial Relations*, 9, 197-214.
- Tannenbaum, A. S. (1974). Systems of formal participation. In G. Strauss et al. (Eds.), *Organizational behavior: Research and issues*. Madison, WI: Industrial Relations Research Association.

## PERSPECTIVES

### The Theory-Practice Relationship in Teacher Education: Need There Be Any?

Peter J. Miller  
and  
Gerald D. Taylor  
*University of Alberta*

Perhaps the most persistent and vexing problem area in teacher education programs is the place occupied by the so-called theoretical studies in education. No one seriously questions whether teacher trainees ought to develop appropriate and adequate subject matter expertise through their university studies. Similarly, no one argues against the need for teacher trainees to be given extensive opportunities to practice teach before they are certified as competent teachers. Subject matter expertise and practical experience—these are the obvious and undisputed elements of teacher education programs. But about the place and purpose assigned to professional theoretical course work and the relationship that this abstract theoretical study of education ought to have to the actual performance of the teacher in the classroom, there exists no such agreement.

In recent years this professional theoretical component of teacher preparation programs has been the subject of considerable criticism and dissatisfaction. Politicians, administrators and students have increasingly demanded an initial preparation which fosters not so much the theoretical abstract understanding of education but rather the specific skills and competencies which beginning teachers are believed to need if they are to perform effectively in classrooms as they presently exist. In response to this demand, institutions of teacher education have

---

Peter Miller (PhD, Alberta) is Dean of Student Services at the University of Alberta and a Professor in the Department of Educational Foundations. His major teaching and research area is the history of childhood as this relates to the contemporary belief systems of parents and teachers.

Gerald Taylor (PhD, Toronto) is an Associate Professor at the University of Alberta in the Department of Educational Foundations. His major research and teaching interests include the sociology of learning, the sociology of teaching, and school-community relations.

tended to center their programs on the most directly practical aspects—instructional and management skills and various competencies developed within an extended practicum framework. With this emphasis on teaching competencies and the practicum, both geared towards entry level certification, the more abstract theoretical courses in teacher education programs have faced considerable pressure, either to show a payoff that is directly related to classroom teaching or to step aside in favor of more skill and practice oriented experiences. Even highly theoretical education courses, which on the surface appear to have little connection with the teaching act (e.g., history, philosophy and sociology of education), have been evaluated in terms of their perceived relevance to teaching performance and classroom effectiveness. Within this general view of teacher education, then, it is held that the relationship between theoretical course work and the instructional and management skills demonstrated in the practicum experience must be a direct one. In effect, university based theoretical course work is expected to affect—that is, to be causally related to—performance in the classroom, both as a student and as a certified teacher. And when it is shown or believed that this has not occurred, the typical response is to seriously question the relevance of such course work and hence its value to initial teacher preparation.

What research has been reported, however, suggests that there is little relationship between performance in theoretical course work and performance in teaching. Jencks (1972, p. 187) conclude from a number of studies relating professionals' performance on the job to college grades that this relationship is much weaker than commonly supposed. In the case of elementary and secondary school teachers, they point out that college grades and observer ratings of teacher performance correlate only between 0.2 and 0.3. Other researchers have uncovered similar statistical trends within the field of teacher education. Emanuel, Larimore, and Sagan (1975), in a study of 42 secondary social studies majors, found that student teaching performance was not significantly correlated with final grades in any of seven different education courses. The correlation coefficients reported in their study ranged from a low of 0.05 between grades in a general teaching methods course and student teaching performance to a high of only 0.28 between the latter and grades in a senior social studies methods course. And although Twa and Greene (1980) recently found, in a study of 139 teacher trainees at the University of Lethbridge, that students' grade point average in their teaching major was significantly correlated with their performance in practice teaching, the magnitude of this correlation was only 0.26 (p. 8). Similarly, while Emerson, Elford and Scaldwell's (1982) study of 112 Althouse teacher trainees found a positive association between practicum success and performance in foundational studies (i.e., theoretical course work in psychology, philosophy and sociology), the phi coefficients ranged from a low of 0.03 in the case of philosophy to a high of only 0.34 in psychology. Further evidence in support of these findings is provided by a recent study at the University of Alberta conducted by Taylor and Miller (1984) in which little correlation was found to exist between grades achieved in professional courses during a professional year and practice teaching ratings.

These findings are in line with those of other researchers, findings which have led Burton to conclude from the literature that performance in teacher training courses bears little relationship to success in teaching.<sup>1</sup> Indeed, the evidence currently available tends to support the observation drawn by Cortis and Satterly (1968) who, in reference to teacher education in England, concluded that there is an apparent dichotomy between the criteria for success within the academic/theoreti-



cal component of teacher education and the criteria needed for success within the practicum component. Good students apparently do not necessarily make good teachers. Understandably, such evidence tends to increase the concern expressed by those who have grown impatient with theory and skeptical of its value to practice; it also seriously weakens the position of those who hold tenaciously but precariously to the view that theory must be the basis on which professionals, particularly teachers, build their practice.

For many people working in the field of education this weak association between performance in theoretical professional courses and performance in teaching will come as no surprise. Students who have completed at least some practice teaching, classroom teachers, and school personnel engaged in the hiring of teachers frequently comment that the qualities, aptitudes and skills required for successful performance as a university student and successful performance as a classroom teacher are sufficiently different that one ought not to expect much of a relationship between the two. However, for those concerned with the design of teacher education programs, this weak relationship is disturbing and of crucial importance since it raises questions about the appropriateness and, indeed, the validity of the current approach to teacher education. It casts serious doubt, in fact, on the assumption that existing structures, mechanisms, and strategies in teacher education can, under even the very best conditions, ever effectively integrate theory and practice.

In itself the well documented lack of any relationship between performance in theory and performance in practice carries with it no implications for teacher education. It is merely an observed and, for most people not so surprising, phenomenon. When, however, this phenomenon is considered in the light of a particular model of teacher education, with all its various value commitments and assumptions about the priorities of professional education, then a number of important conclusions can be drawn and, more importantly, recommendations formulated. Consider in this connection the significance of this observed lack of a theory-practice performance relationship in the light of three models of teacher preparation.

A weak association between the two modes or domains of performance is unsettling to all those who believe that university-based initial teacher education, whether it culminates in a degree or not, should focus primarily on the development of the entry level teaching skills needed for effective performance in a conventional classroom. But it is most disturbing to those who insist that it is the teaching practice or practicum experience that is of fundamental importance in the development and firming up of these necessary teaching competencies. For this practice-driven or competency-based model, the practicum component of initial teacher education is, of course, central to the entire program. The Bachelor of Education degree program's professional course work thus tends to be conceived, designed and evaluated in light of the skills and competencies needed by teacher trainees to successfully complete practice teaching and thus become eligible for initial certification as beginning classroom teachers.

For those committed to this practice-driven model of a Bachelor of Education degree program, when little apparent association exists between being a good student and being a good teacher, one of two related conclusions can be drawn. First, the theoretical courses may be poorly designed or out of touch with the realities of classroom teaching, and hence stand in need of basic change in content or mode of delivery. More fundamentally, it may be concluded that conventional formal professional course work has lost its validity or *raison d'être*. The implications of such a position are, as Emanuel, Larimore and Sagan (1975) point out, that

Colleges of Education would be well advised to examine closely their teacher training programs and decide whether all of these courses are worth keeping in the program or whether their content should be integrated with actual classroom experiences in order to establish a greater relationship between theory and practice. (p. 248)

From the standpoint of this model, then, the theory component must be more closely molded around the practicum since the latter operates within the real world of the classroom. And this can be accomplished most effectively, Emanuel, Larimore, and Sagan (1975) inform us, by replacing conventional theory courses with what they refer to as "experientially-based" courses (p. 248). These, they argue, will integrate the objectives of the courses contained in teacher preparation programs, demonstrate to the trainees the utility of theory to practice, and strengthen the relationship between course performance and performance in student teaching. These benefits will result because under such conditions the two behaviors being compared become very similar.

Advocates of this model hold, therefore, that the theory component must be re-fashioned so that the competencies needed and the criteria of evaluation in theoretical studies converge with those of the practicum. Reform of teacher education is envisaged, in effect, primarily in terms of redesigning the theoretical elements since it is their relevance, utility and hence legitimacy that are in question. Conversely, from this perspective, improvements of the practice component entail, for the most part, an extension of the practicum rather than any alteration of its present nature and purpose.

But to be committed to an initial certification model of teacher education and thus of the Bachelor of Education program, it is not also necessary to be committed to a practicum-driven or competency-based model. Consider, for example, a model of teacher preparation which runs counter to it and as such leads to substantially different conclusions about the weak association between being a good student and being a good teacher. This alternative perspective, which may be termed the theory-driven model, holds that the practicum experience is not the most important element of the teacher education program; it should not, therefore, dictate the nature and purposes of theoretical studies during the initial preparation of teacher candidates. On the contrary, exponents of this perspective argue that sound practice relies heavily on a strong foundation not only of general knowledge derived from liberal arts but also on in-depth understandings of teaching and learning that can be fostered by professional education courses. Hence, teacher trainees must engage extensively in the study of "the most recent and well-founded knowledge in the disciplines of educational study" (Taylor, 1981, p. 3). They must also reflect on the nature of this knowledge and its implications for learning and teaching. And no less significantly, they must demonstrate an awareness, understanding and appreciation of these theoretical underpinnings of sound teaching practice. The importance of such understanding cannot be overstated since it is this body of educational theory that provides the basis on which to build, in Joyce's (1981) terms, "the intellectual and interpersonal tools to continue to study and expand the range of highly developed skills" (p. 2) required of teacher candidates if they are to be permanently committed, as Taylor (1981) puts it, "to personal learning and professional development" (p. 3).

From the perspective of this theory-driven model of initial preparation, the practicum is still of central importance; its primary purpose, however, is to serve as a bridge between educational theory and practice. As expressed in the Government of Alberta's Theory to Practice report (1981), practicum experience should be de-



signed to “bring theoretical and practical aspects of undergraduate education into a clear and interacting type of relationship and to aid students in applying what they learn theoretically to simulated and actual situations” (p. 14). In effect, the practicum is envisaged as a kind of laboratory experience in which the theory studied in course work can be, in some sense, tested in the classroom.

But if the practicum is to help foster among teacher candidates an appreciation of the relevance and utility of theory to actual classroom teaching and, moreover, culminate in an ability to translate theoretical concepts into teaching practice, it must be designed specifically with these objectives in mind. In particular, to again borrow from the Government of Alberta’s Theory to Practice report (1981), the practicum must provide ample opportunity for student trainees to “create, innovate, test and revise theories and methodologies” (p. 15). Moreover, the analysis of practice within it must be, according to Taylor (1981, p. 10), “systematically conducted in terms of concepts and theories which feature in other areas of the professional sequence” comprising the teacher education program.

However, according to proponents of this theory-driven model, it is highly doubtful that any clear articulation between theory and practice can be achieved within a program of teacher education which is firmly tied to the certification of entry level teaching competencies and skills—that is, when it is practice-driven. According to Joyce (1981, pp. 1-2), initial preparation should prepare teacher candidates:

1. For the lifelong study of the world, the self and academic knowledge.
2. The lifelong study of teaching.
3. To participate in school renewal efforts including the creation and implementation of innovations.
4. To approach the generic problems of the workplace (the school and the classroom).

Such objectives are extremely difficult to achieve, however, when conventional practicum experiences have as their goal socializing teacher candidates to the role of teachers in conventional classrooms. For as Joyce notes, the student teaching experience tends to introduce and initiate prospective teachers into the conventional wisdom, procedures and values of the current approach to practice rather than systematically to provide for the opportunity to test out and implement emerging or more progressive strategies suggested in theoretical course work. Moreover, during the practicum many teacher candidates also develop an attitude of pragmatism which turns them “from university and gradually alienates them from the intellectual roots of their profession” (p. 12). They tend, therefore, to “become concerned with quick and easy solutions that will make immediate differences to their classroom rather than the lifelong study of the more complex and interesting ways of approaching children” (p. 12); and they become “more constrained in their behavior, more authoritarian and more resistant to change” (p. 9). Operating under the pressure of summative credential-oriented evaluation, they come to “feel tremendous stress . . . [which] has the effect of driving the[ir] self-concept . . . toward (in Maslow’s terms) a survival orientation” (p. 9).

For proponents of this theory-driven model, therefore, if the abilities, skills, and competencies required for success in theoretical studies are dramatically different from those needed for successful performance in the conventional practice, what stands in need of serious revision are the latter, not the former. In other words, the evidence indicating that good students do not necessarily make good teachers is in-



interpreted to mean not that theoretical course work is irrelevant to teaching effectiveness in the practicum experience, but rather that the practicum, and indeed, school practice itself, are being conducted and evaluated inappropriately.

From the lack of any strong relationship between performance in theoretical course work and performance in practice can be drawn, then, two widely differing sets of implications for teacher education, the one resulting in a critique and proposed revision of the theory element, the other in a critique and proposed revision of contemporary practicums, and indeed, of classroom practices. Both the practicum-driven and the theory-driven models, however, are similar in that they are both based upon a conviction that theory and practice should be both theoretically and empirically related. Both insist, therefore, that entry level certification and the Bachelor of Education degree should be granted upon successful completion of a program in which professional theoretical course work has been, in some fashion, causally related to the practicum or student teaching experience. In this respect both models are based on convictions and assumptions about professional education which, according to the evidence at least, have little basis in reality. Nor is it clear if a knowledge of theory, that is, of the various concepts and ideas dealt with in professional course work, can be even logically related to the various competencies, skills and attitudes which are required for successful performance in classrooms as they presently exist.

There are thus good grounds for suggesting that faculties of education should consider an alternative model of professional education, one that stresses not the logical and/or empirical interdependence and interrelationship of theoretical studies and certification-oriented practical training but rather their discreteness or separateness. Each may well be of fundamental importance in professional training but this does not mean that they are therefore causally related. Indeed, each should be seen as having its own rationale and each its own unique and appropriate criteria for evaluating success or failure.

It is important in this connection to separate out what can be achieved during a university degree program from what can be achieved in an apprenticeship or internship setting. What a Bachelor of Education degree program provides is knowledge. To begin with, and most importantly, it provides a knowledge of the subject matter to be taught; and about the value of this kind of knowledge there is no debate. But it also typically includes elements of theoretical professional studies which provide the basis of an informed understanding of the professional activity in all its contexts. The professional payoff of such knowledge is considerable. To begin with, in the western world this kind of understanding is generally held to be the distinguishing characteristic of a professional activity as opposed to a craft or trade. It is thus closely tied to both income level and occupational status. It is also an indispensable prerequisite for advanced study and research in a professional field and the basis of lifelong learning to which teachers should be committed. Finally, it is directly related to an ability to critically evaluate professional practice, not only within the narrow terms of professional concerns but also within the larger social context; it is thus linked to a willingness to experiment and innovate. It should be noted that this kind of knowledge and the attitudes and values it embodies exist quite independently of the competencies needed for success in a conventional classroom.

The criteria for demonstrating such knowledge in a university setting are well known and need no further elaboration. They are certainly clearly different from those which govern the evaluation of performance in a practicum activity which has

as its primary goals the development of needed job competencies and, equally important, socialization of the trainee into the professional or conventional order of things. To allow these two discrete areas of professional education to spill over into each other and to exercise a strong influence on each other is to compromise the integrity of both. On the one hand, it seriously weakens the quality of what is essentially an academic activity while, on the other, it intellectualizes unnecessarily what is basically a practical and socializing experience.

The implications of such conclusions for the professional education of teachers are profound. The Bachelor of Education degree should be freed from its current preoccupation with entry level skills and become an academic credential awarded on the basis of academic considerations. The theoretical or professional studies component of the degree program would, of course, involve students in a variety of experiences that would take place in school settings. These laboratory experiences, however, would not be viewed as preparation or training for effective functioning in a conventional classroom—that is, as a method or means of acquiring a particular repertoire of specific teaching competencies. Rather, the nature and function of these experiences would be determined by the academic requirements of theoretical course work.

Whether school boards and teachers would be willing to throw open their classrooms to the experimental activities and inevitably critical gaze of university instructors and their students is doubtful. And, of course, it is not clear whether this kind of use of conventional classrooms could be justified or, indeed, whether it would be worthwhile. It may be that faculties of education should create their own laboratory schools in which students could observe a variety of classroom phenomena, critically evaluate the appropriateness of particular kinds of procedures, and experiment in all areas of teaching without any concern that their efforts will be judged on grounds that have to do with conventional norms of teaching effectiveness.

The Bachelor of Education degree, when conferred, would certify that the holder had achieved a certain level of general knowledge, subject matter expertise and a theoretical understanding of the various aspects of education. It would thus be similar in kind to the degrees awarded in other professional fields, notably business, law and medicine. Its possession would permit the holder to apply for a further program of professional training, an internship, to be conducted in a school setting. When successfully completed, the internship would lead to certification as a professional teacher. The details of such internship programs need not concern us here, although they would probably have a good deal in common with existing conventional extended practicum programs. They would have as their goal the training of the intern in all the needed job competencies required for effective functioning in the contemporary classroom. The internship training program could be developed on the assumption that the interns were familiar with the theoretical aspects of education. It would not, however, need to be designed in such a way that its relevance or practicality would be compromised by importing into it any abstract theoretical distractions.

Whether this separation of theory and practice—of academic and non-academic elements—in teacher education will produce better teachers, we have no way of knowing. But it seems likely that such programs will produce teachers who will be better educated during their degree program, and better trained during their internship period. And these are surely steps in the right direction.



*Notes*

1. Cited in Emanuel, Larimore and Sagan (1975, p. 246).

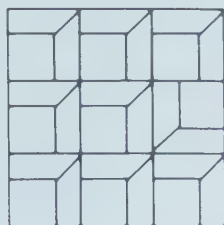
*References*

- Cortis, G. A., & Satterly, D. J. (1968). Personality profiles of student teachers. *British Journal of Psychology*, 32(2), 115-122.
- Emanuel, J., Larimore, D., & Sagan, E. (1975). The relation of selected academic program variables to student teaching performance. *Contemporary Education*, 46(4), 245-248.
- Emerson, G., Elford, G., & Scaldwell, W. (1982). Who succeeds in practice teaching. *News Quarterly*, 1(5), 1-5.
- Faculty of Education. (1983). *Handbook for practicum programs in the elementary schools, 1983-84*. University of Alberta.
- Government of Alberta. (1981). *Theory to practice: Report of the committee to evaluate the extended practicum at Alberta universities*.
- Jencks, C. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Harper & Row.
- Joyce, B. R. (1981). The phoenix plan: Essential reform in teacher education. In Government of Alberta, *Theory to practice: Report of the committee to evaluate the extended practicum at Alberta universities*. Appendix B.
- Taylor, G. D., & Miller, P. J. (in press). Professional course work and the practicum: Do good students make good teachers? *Canadian Journal of Education*.
- Taylor, W. (1981). Paper for the committee on the evaluation of teacher education programs in Alberta. In Government of Alberta, *Theory to practice: Report of the committee to evaluate the extended practicum at Alberta universities*. Appendix A.
- Twa, R. J., & Greene, M. (1980). Prediction of success in student teaching as a criterion for selection in teacher education programs. *The Alberta Journal of Educational Research*, 26(1), 1-13.



## Table of Contents, Volume XXX, 1984

- Allison, D. E. *The effect of homogeneous vs. heterogeneous matching-item format on test performance and reliability.* No. 2, June, 91-94.
- Belanger, J., and Martin, R. G. *The influence of improved reading skill on writing skill.* No. 3, September, 194-212.
- Carey, S. T., and Cummins, J. *Communication skills in immersion programs.* No. 4, December, 270-283.
- Dash, U., and Maguire, T. *A detailed analysis of group differences on the California short-form test of mental maturity between 1956 and 1977.* No. 2, June, 95-103.
- Evans, P., Mickelson, N., and Smith, N. C. *Computer-assisted instruction in college composition.* No. 4, December, 311-319.
- Friesen, D., Holdaway, E. A., and Rice, A. W. *Factors contributing to the job satisfaction of school principals.* No. 3, September, 157-170.
- Gayle, G. M. H. *Second language teaching styles: Making or breaking attitudes.* No. 1, March, 49-67.
- Gunderson, L. *One last word list.* No. 4, December, 259-269.
- Houle, R., and Montmarquette, C. *An empirical analysis of loans by school libraries.* No. 2, June, 104-114.
- Hoy, W. K., and Sousa, D. A. *Delegation: The neglected aspect of participation in decision making.* No. 4, December, 320-331.
- Levin, B. *Development and initial results of a longitudinal secondary follow-up study.* No. 1, March, 1-13.
- Maynes, F. J., and Ross, J. A. *Integrating school curricula with social concerns.* No. 3, September, 179-193.
- Merritt, F. M., and McCallum, R. S. *The relationship between simultaneous-successive processing and academic achievement.* No. 2, June, 126-132.
- Miller, P. J., and Taylor, G. D. *The theory-practice relationship in teacher education: Need there be any? (Perspectives)* No. 4, December, 332-339.
- Mueller, H. H., Dash, U. N., Matheson, D. W., and Short, R. H. *WISC-R subtest patterning of below average, average, and above average IQ children: A meta-analysis.* No. 1, March, 68-85.
- Mueller, H. H., Mancini, G. J., and Short, R. H. *An evaluation of the diagnostic efficiency of the Wechsler Intelligence Scale for Children—Revised.* No. 4, December, 299-310.
- Nixon, M. T., and Bumbarger, C. S. *Prospective teachers: Views of programs and needs.* No. 3, September, 226-237.
- Pansegrau, M. *Teachers' perspectives on inservice education.* No. 4, December, 239-258.
- Rampaul, W. E., Singh, M., and Didyk, J. *The relationship between academic achievement, self-concept, creativity, and teacher expectations among Native children in a northern Manitoba school.* No. 3, September, 213-225.
- Robinson, N., and Stacey, C. *Serving on the school board: A political apprenticeship for higher office.* No. 2, June, 115-125.
- Staab, C. F. *Student teachers' perceptions of their competence versus the perceptions of university supervisors and sponsor teachers.* No. 4, December, 284-298.
- Walberg, H. J., Singh, R. P., and Tsai, S. L. *The achievement productivity of psychological variables: A replication and extension in Rajasthan, India.* No. 3, September, 171-178.
- Westbury, M. *The link between universities and employment in Alberta.* No. 1, March, 38-48.
- Whyte, L. A. *Characteristics of learning disabilities persisting into adolescence.* No. 1, March, 14-25.
- Wideen, M. F., and Hopkins, D. *Supervising student teachers: A means of professional renewal?* No. 1, March, 26-37.



# Journal of Education

Published by  
Boston University  
School of Education

Winter 1983 Volume 165, Number 1

## **Literacy and Ideology**

focusing on the theoretical relationship of reading, literacy, and ideology with specific studies of the French immersion program in Canada and the literacy programs in the Cape Verde Islands and Guinea-Bissau. Papers by

**Paulo Freire**  
**Henry Giroux**  
**James Donald**  
**Paul Olson**  
**Donaldo Macedo**

Spring 1983 Volume 165, Number 2

## **The Dialectics of Evaluation**

focuses on the social and political goals of educational environments and their implications in evaluation and testing: Studies of the pre-school environment, student classification, and testing considered in the light of gender and social policy.

Winter 1984 Volume 166 Number 1

## **Class, Gender, and Schooling**

Papers by

**Jean Anyon**  
**Madeline Arnot**  
**Miriam David**  
**Rosemary Deem**  
**Jane Gaskell**  
**Ann Manicom**

Requests for single copies or subscriptions to:

**Journal of Education**  
**Boston University**  
**605 Commonwealth Avenue**  
**Boston, Massachusetts 02215**

|                                      |         |
|--------------------------------------|---------|
| Single issue:                        | \$ 8.00 |
| Individual one year subscription:    | \$16.00 |
| Institutional one year subscription: | \$21.00 |





















## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in triplicate and are not to exceed 25 pages in length.
2. An abstract of approximately 100 words in length, typed on a separate page, is to be included.
3. The author's name and affiliation should appear only on a separate title page to ensure anonymity in the reviewing process; the title of the paper should appear on the first page of the manuscript as a means of identification.
4. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publication Manual.
5. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
6. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
7. Sources should appear in parentheses following the reference, citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by authors' last names at the end of the manuscript under the heading, *References*.
8. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 7 above.
9. Spelling shall conform to *Webster's New Collegiate Dictionary*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
10. In matters of style, the third edition of the APA Publication Manual is considered definitive.

